Continuous-time Markov decision processes: theory, approximations and applications

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of Doctor in Philosophy
by

Yi Zhang

Supervisors: Dr. Alexey Piunovskiy and Dr. Damian Clancy

June 2010

Abstract

In a nutshell, this thesis studies the solvability issue and the approximation issue of CTMDPs (Continuous-Time Markov Decision Processes).

In Chapter 2, under standard conditions allowing unbounded (both from above and below) transition rates and cost rates, for a discounted CTMDP in a Borel state space, when there is only one performance criterion, we develop the dynamic programming approach and establish the existence of a deterministic stationary optimal policy. When there are more than one criteria, we develop the convex analytic approach and show the existence of a (possibly randomized) stationary optimal policy.

Chapters 3, 4 and 5 address the approximation issue of CTMDPs by studying the accuracy of the fluid approximations for three specific CTMDPs. In Chapter 3, for a Birth-and-Death process with an absorbing state, we firstly propose an appropriate fluid model, which can differ from the naturally looking one, as confirmed by means of an example. Then, for this fluid model we show that a feedback translation of the fluid optimal policy results in a satisfactory policy, which achieves some optimality asymptotically, with the rate of convergence (also called its efficiency) being estimated explicitly involving only the primitives. The same issues of fluid approximations are investigated in Chapter 4 for EOQ (Economic Order Quantity) and EPQ (Economic Production Quantity) models. The translations in Chapters 3 and 4 are both of a feedback type, and their efficiencies are both $O(\frac{1}{n})$, where n is the fluid scaling parameter. In Chapter 5, we consider the tracking policy which is resulted in by a time-dependent but nearly state-independent translation of the fluid optimal policy. For a bandwidth-sharing network over a finite horizon, firstly, without extra conditions being imposed on the network parameters, we show the efficiency of the tracking policy to be $O(\frac{1}{\sqrt{n}})$; and secondly, after imposing some extra conditions on the network parameters, we show that the tracking policy could be also of the efficiency $O(\frac{1}{n})$. By means of an example, we are assured that the feedback translation considered in Chapters 3 and 4 cannot generally result in a policy (for the stochastic model) of the efficiency better than $O(\frac{1}{n})$.

In Chapter 6 we investigate a fluid model of an Internet router under the MIMD (Multiplicative Increase Multiplicative Decrease) congestion control. We firstly study

the long run behaviour of the trajectories (the number of packets in the router along the time), for which a limiting regime, in the form of a cycle, is shown. After that, with in mind two performance criteria, we then study the optimal buffer size, given as a Pareto set. Finally, when the parameters are tuned to correspond to STCP (Scalable TCP), the obtained results are verified by simulations.

To Daniel Stephen Morrison, a very well liked friend who has a decisively important positive influence on my attitude toward life.

Contents

Acknowledgement

It is a pleasure to thank those who made this thesis possible. Especially, I would like to express my wholehearted gratitude to my primary supervisor, Dr. Alexey Piunovskiy (University of Liverpool, UK), whose teaching, supervision, patience and encouragement will be greatly valued forever. During the time of preparation for this thesis, extremely helpful suggestions have been received from for example, Prof. Anyue Chen (University of Liverpool, UK), Dr. Damian Clancy (University of Liverpool, UK) and Dr. Toby Hall (University of Liverpool, UK). On various occasions during his visit to Liverpool in 2008, Prof. Xianping Guo (Zhongshan University, China) also advised me useful things. In 2010, Dr. Maaike Verloop (BCAM, Spain) kindly presented me her Ph.D thesis for reference. Most of the materials presented in this thesis are based on the research done in collaboration with Dr. Alexey Piunovskiy, Dr. Konstantin Avrachenkov (INRIA-Sophia Antipolis, France) and Dr. Urtzi Ayesta (LAAS-CNRS, France). In particular, Dr. Konstantin Avrachenkov kindly hosted and supervised me at INRIA-Sophia Antipolis in 2008 and 2009 for six months.

There is no doubt that I am always indebted to my parents for their consistent support. Similar words should be applied to my friends, too.

Finally, I am grateful to the financial support received in forms of ORSAS, University of Liverpool Graduate Association (Hong Kong) scholarships, and standard research studentships at the University of Liverpool.

Chapter 1

General introduction

This thesis concerns about CTMDPs, where the system dynamics can be modelled as jump processes, and one (the so called decision maker) can implement some control on the jump (transition) rates continuously in time over a given time horizon, in order to achieve the optimality regarding some predetermined performance criterion (criteria). Therefore, the ultimate aim in CTMDPs is to find (provided its existence) the optimal decision rule (policy) out of all the feasible ones. There are many ways to classify CT-MDPs. For example, depending on the number of performance criteria kept in mind by the decision maker, CTMDPs are called unconstrained (if there is only one criterion) or constrained (if there are more than one). An alternative classification of CTMDPs is according to the underlying horizon, which can be finite or infinite. In case of a finite horizon, typically, the criterion (criteria) may take the form of minimizing the expected total cost, whereas if the horizon is infinite, one often aims at minimizing one of the following: the expected total cost up to when the system reaches a certain point, the expected total discounted cost, or the long run average expected cost. Hence, this raises yet another classification of CTMDPs, which is by the exact forms of the optimization criteria. In fact, the aforementioned criteria are often referred to as "basic" (see Hernández-Lerma and Lasserre (1996)), and we refer readers to Guo and Hernández-Lerma (2009) for other "advanced" ones. Here we confess that all the terminologies as well as some mathematical expressions used below in this chapter are by no means rigorous: they are only to give preliminary impressions, and their accurate definitions will be given in later chapters.

Standard examples of CTMDPs include controlled queueing systems, inventories and BS (bandwidth-sharing) networks.

Preliminary example (a): Consider an $M/M/1$ queueing system (which, if one likes, can be also read as a Birth-and-Death process), where the instantaneous state Y_t takes values from the state space $S = \{0, 1, 2, \dots\}$. Suppose one can control the arrival rate as well as the service rate, so that for any fixed $x \in S$, the jump rate $q(y|x,a)$ concentrated on x, $x+1$, $x-1$ (if $x=0$, then we do not consider $q(x-1|x,a)$) is also a-dependent, where $a \in A(x) \subseteq A$ stands for the chosen action (control), A is the action space, and $A(x)$ is the admissible action space given the current state x. Suppose the cost rate takes the form of $c(x,a)$, and one aims at minimizing the expected total cost $E^{\pi}_{Y_0}\left[\int_0^\infty c(Y_t,a(t))dt\right]$ over all feasible π, where π and Y_0 stand for the policy and the initial state. By the way, this expectation is called the performance functional of the policy π (for the concerned performance criterion), and the infimum of performance functionals over all feasible π is called the value function of the CTMDP. Moreover, for our particular interest in the system before it firstly goes empty as well as technical reasons, we typically put $q(y|0,a)=c(0,a)=0$. That is, state zero is absorbing. Therefore, the underlying CTMDP is called an absorbing one. Sometimes, it is desirable not to take any state as absorbing. On such occasions, for the sake of technicalities as well as the concept of present values, one often instead aims at minimizing the expected total discounted cost $E^{\pi}_{Y_0}\left[\int_0^\infty e^{-\alpha t}c(Y_t,a(t))dt\right]$, where $\alpha>0$ is a fixed discount factor. Then the corresponding CTMDP is called a discounted one. If in the mean time of minimizing $E^{\pi}_{Y_0}\left[\int_0^\infty e^{-\alpha t}c(Y_t,a(t))dt\right]$, one must ensure $E^{\pi}_{Y_0}\left[\int_0^\infty e^{-\alpha t}c_1(Y_t,a(t))dt\right] \leq d_1$, where $c_1(x,a)$ is another cost rate, and d_1 is a fixed constant, then the CTMDP is called a constrained discounted one.

Preliminary example (b): Consider an EOQ model, where the instantaneous state Y_t stands for the inventory level. Let the state space, action space and admissible action space be $S=\{0,1,2,\ldots\}$ and $A=A(x)=\{1,2,\ldots\}$, respectively. Consider the jump rate $q(y|x,a)=q(y|x)$ concentrated on x and $x-1$ if $x>0$, and $q(y|0,a)$ concentrated on a and 0. In words, the inventory level decreases gradually like a pure Death process up to state zero, at which the decision maker will order new inventories, which will arrive after some exponentially distributed delay. The order quantity is controlled so as to minimize $\lim_{T\to\infty}\frac{1}{T}E^{\Phi}_{Y_0}\left[\int_0^T c(Y_t,\Phi(Y_t))\right]$ out of all functions Φ from S to A. Then the underlying CTMDP is one with a long run average cost. Here $\Phi(0)$ indeed stands for the order quantity, and in the context of Inventory theory, its corresponding performance functional (the long run average expected cost) is called total cost per unit time, or in short TCU. If a policy is optimal, then the corresponding order quantity is called "economic". By the way, policies taking the form of a function Φ are called deterministic stationary.

Preliminary example (c): Consider a simple linear bandwidth-sharing network (see Massoulié and Roberts (2000); Verloop (2009)) with two resources and three types of flows, where the instantaneous state is $Y_t=(Y_t^1,Y_t^2,Y_t^3)'$, with Y_t^k, $k=1,2,3$ standing for the number of flows of type k. Here the symbol of $'$ means the transposition. Let the state space be $S=\{(x_1,x_2,x_3)'\}\subseteq(0,1,2,\ldots)^3$, and consider the action space as well as the admissible action space $A=A(x)=\{(a_1,a_2,a_3)'\}\subseteq(\mathbb{R}_0^+)^3$ such that $R(a_1,a_2,a_3)'\leq Z$ with $R=\begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}$ and $Z=(z_1,z_2)'\in(\mathbb{R}_+^0)^2$. Here $\mathbb{R}_+^0$ (resp.

a_k, $k = 1,2,3$) stands for the set of nonnegative real numbers (resp. the allocated capacity to flows of types $1,2,3$), and Z indicates the capacity of each resource. Note, in a bandwidth-sharing network, if a flow is routed on several resources, then each of those resources must provide the same capacity to the flow. The transition rate $q(y|x,a)$ is concentrated on x_k+1, x and x_k-1, $k=1,2,3$, where if $x_k=0$, then we do not consider x_k-1. With the cost rate $c(x,a)$ and a fixed horizon $0<T<\infty$, one aims at minimizing $E^{\pi}_{Y_0}\left[\int_0^T c(Y_t,a(t))\,dt\right]$. The underlying CTMDP is usually called a finite one.

Note that in Preliminary example (b), the CTMDP is restricted to the class of deterministic stationary policies. There are mainly two reasons for making such primary assumptions. One is that in practice sometimes the system operator is subject to technical difficulties and can only implement such simple policies. The other one is solely for simplicity. As far as the latter reason is concerned, the general theory of CTMDPs has been developed in order to cover the following issues:

Problem set (a): Does it add values by allowing into consideration more complicated policies, which for example, can be randomized as well as dependent on the past information (history) about the system? In other words, is the class of deterministic stationary policies sufficient for solving the underlying optimization problems? In case the answer is positive, does there exist a deterministic stationary optimal policy?

On the other hand, suppose an exact optimization problem in the form of a CTMDP with local transitions (like the $M/M/1$ queueing system in Preliminary example (a)) is fixed together with its parameters (the jump rate and so on). Then one must solve for its optimal policy. Primarily, it is desirable to solve the CTMDP analytically with its optimal policy as well as its value function, which, when viewed as a pair, is often referred to as the optimal solution. Unfortunately, this is well known to be a difficult task, and only relatively few examples, where CTMDPs are solved explicitly have been available in the literature, see for instance, Kitaev and Rykov (1995); Tan et al (2010); Verloop (2009) and (Piunovskiy 1998, Sec.7). Therefore, algorithms such as value iterations and policy iterations (see Guo and Hernández-Lerma (2009)) have been developed to obtain the optimal solution numerically. Unfortunately, when the state space and action space are large, as they typically are, such algorithms become much less efficient. Consequently, one gets satisfied with obtaining policies, whose performances are close to that of the optimal policy. Intuitively, this may be achieved by solving a problem close (in some sense) to the original one, and this problem must be easier for investigations. Having in mind that the main difficulty for solving a CTMDP can come from its discrete and stochastic behaviour, a natural candidate problem to do the job would be its continuous and deterministic analogue, the so called fluid model[1].

Preliminary example (d): For the EOQ model described in Preliminary example (b),

[1]Our introduction of the term "fluid model" is in a strong taste of engineers. A different definition (which we do not consider in this thesis) of that term states it to be the collection of all the fluid limits, though the term "fluid limit" in its own right varies with different authors, see Foss and Kovalevskii (1999) and the references therein.

let us put $q(x-1|x,a) = \mu > 0, \forall\, x > 0$. Then its fluid model looks like the following: denoting $y(t) \geq 0$ the instantaneous state and φ, which maps $[0,\infty)$ to $(0,\infty)$, any fixed policy (for the fluid model), then with the initial condition $y(0) = y_0 = Y_0$, the system dynamic of the fluid model is governed by $\frac{dy}{dt} = -\mu \ \forall\, y > 0$, and $y(t^*+) = \varphi(0)$, where t^*+ stands for the moment immediately after $\{y(t),\ t \geq 0\}$ reaches state zero. Here we choose not to take into account the delay about the ordering, because in practice keeping the inventory empty is not desirable, and in the fluid model everything is deterministic, meaning that one can always effectively get rid of any delay by ordering beforehand. Then one aims at minimizing $\lim_{T\to\infty} \frac{1}{T} \int_0^T c(y, \varphi(y))dt$, where we have put y instead of $y(t)$ for brevity.

Indeed, fluid models are generally easier to tackle (see Gajrat and Hordijk (2005)), and they can be solved, while the analytical study of the original stochastic ones look untractable (see Verloop and Núñez Queija (2007); Verloop (2009)). Calling approximating the original CTMDP using its fluid model "fluid approximations", the following issues immediately arise:

Problem set (b): Are fluid approximations accurate, at least in some sense? In addition, given a policy for the fluid model which is then translated in some way to one for the stochastic model (in the form of a CTMDP), how different is the performance functional of this translated policy in the stochastic model from that of the (pre-translated) policy in the fluid model? Here we are mainly concerned with CTMDPs with local transitions.

Note, in Problem set (b) we emphasize that one is interested in the accuracy of fluid approximations "in some sense". As a matter of fact, in the general sense it is very easy to encounter situations where the fluid model fails to provide reasonable approximations. For example, suppose we are concerned with a controlled $M/M/1$ queueing system (recall Preliminary example (a)), where the initial state is $Y_0 = 1$, $q(x+1|x,a) = \lambda > 0$ and $q(x-1|x,a) = \mu - a$ with $a \in A(x) = A = \{0, \mu\}$, $\mu > 0$ and $\lambda > \mu$ (by a small difference). We shall fix the absorbing state to be some big enough integer, so that the state space S is indeed finite. Suppose the cost rate is $c(x,a) = CI\{x=0\} + aI\{x=1\}$, where $I\{\cdot\}$ is the indicator function and $C > 0$ is a sufficiently big (in relation to μ) penalty incurred whenever the system reaches state zero. Being restricted to the class of deterministic stationary policies, we aim at minimizing $E_1^{\Phi}\left[\int_0^\infty c(Y_t, \Phi(Y_t))dt\right]$, which can be seen well defined for any Φ mapping S to A. Clearly, since C is much bigger than μ, which is then rather close to λ, the optimal control at state one should be $\Phi(1) = \mu$, meaning that the value of its performance functional will be strictly positive. On the other hand, in its fluid model (written down similarly as in Preliminary example (d)), since $y(0) = 1$ and $\lambda - \mu > 0$, one can take $\varphi(y) = 0$, whose performance functional will be zero. Here $\{\varphi\}$ stands for the class of deterministic stationary policies for the fluid model. Therefore, we see that in this example both the optimal policies and the values of performance functionals differ significantly between the stochastic model and its fluid model (see also (Piunovskiy and Clancy 2008, Sec.5, p.427) for a related observation, where the authors considered an epidemic model proposed by Gleissner (1988)). However, if for the stochastic model one speeds up the time as well as scales down the space both linearly with some scaling parameter $n \in \mathbb{N}$ (if $n = 1$, the stochastic model is not scaled), then hopefully in the limiting regime (for large n), the nth stochastic model is well approximated by the fluid model. By the way this scaling (resp. this parameter n) will be called a standard fluid scaling (resp. fluid scaling parameter), whose intuitive meaning will be described in later chapters. Suppose n is fixed. Given a policy for the fluid model, suppose we have translated it according to some way into one for the nth stochastic model. Then the accuracy of this fluid model (approximation) is given by the absolute difference between the values of the corresponding two performance functionals[2], which is generally n-dependent and thus denoted by $\hat{\varepsilon}(n)$. If $\lim_{n\to\infty} \hat{\varepsilon}(n) = 0$, then we say the fluid model is accurate in the sense of fluid scaling, and for the underlying translation mechanism, we call $\hat{\varepsilon}(n)$ its efficiency. In this thesis it is in the sense of fluid scaling that we study the accuracy of fluid approximations. Now the "some sense" in Problem set (b) is clear.

Should the questions in Problem set (b) be answered positively, then one has preliminary beliefs in fluid approximations to (stochastic) optimization problems with local transitions. Such problems are typical in telecommunications. For example, in the Internet most information packets are sent from a source to a destination via basic Drop Tail router(s), which has some bounded space (the so called buffer) to hold a certain amount of information packets when its outgoing link is busy, and reject all the newly arrived ones when the buffer is full. Therefore, the traffic at the router can be viewed as a queueing system. In basic Drop Tail routers, apart from the router capacity, the buffer size is the only parameter to be tuned. Indeed, it is also one of the few parameters that can be tuned by the operators of a TCP/IP (Transmission Control Protocol/ Internet Protocol) network, which makes the choice of the buffer size very important for network designs. Since the traffic in the Internet is of quite complicated stochastic nature, its fluid model is usually formulated for analytical studies (see for example, Avrachenkov et al (2005); Bohacek et al (2003); Bolot and Shankar (1992); Hespanha et al (2001); Khoury and Altman (2004)). Then the following question is of our interest:

Problem set (c): Under the MIMD congestion control scheme (see Altman et al (2005)) which is particularly relevant to Scalable TCP (see Kelly (2003); Khoury and Altman (2004)) as well as the Slow Start phase of TCP New Reno (see Allman et al (1999); Fall and Floyd (1996)), how big should the buffer be sized, if one has in mind two performance criteria of delay (to be minimized) and throughput (to be maximized)?

[2]We emphasize that in the context of optimizations normally it is the performance functional that receives the greatest interest.

Revolving about the aforementioned three problem sets, the main contributions of this thesis are three-folded. Firstly, for constrained discounted CTMDPs with Borel state space and unbounded transition rates, taking into account the class of history-dependent and possibly randomized policies, we answer questions raised in Problem set (a). Indeed, Problem set (a) has been a core topic in the studies of CTMDPs, and has attracted considerable attention from the research community, see Yushkevich (1977); Guo and Dai (2002); Guo and Hernández-Lerma (2003a,b,c, 2009); Guo (2007); Guo and Zhu (2002); Guo and Piunovskiy (2010); Feinberg (2004); Kitaev (1986); Kitaev and Rykov (1995); Piunovskiy (1998, 2005); Yan et al (2008) (this is far from being an exhaustive list, and we refer readers to the most recent survey Guo et al (2006) for more references). However, they are all subject to some limitations: bounded rates are assumed in Piunovskiy (1998, 2005); Feinberg (2004); Yushkevich (1977); Kitaev (1986); Kitaev and Rykov (1995), at most the class of (possibly randomized) policies depending on the current state and time (Markov policies) are allowed for consideration in Guo and Dai (2002); Guo and Hernández-Lerma (2003a,b,c, 2009); Guo (2007); Guo and Zhu (2002); Yan et al (2008), and the state space is required to be countable in Guo and Piunovskiy (2010). Therefore, to our best knowledge, the exposition of constrained discounted CTMDPs in this thesis is based on the most general setting in the current literature. Secondly, we address Problem set (b) and obtain the accuracy of fluid approximations (in explicit formulae involving only primary date) for controlled $M/M/1$ queueing systems, EOQ (as well as EPQ) models and bandwidth-sharing networks, which have been illustrated in Preliminary examples (a, b, c). Formal justifications of fluid approximations, or in other words, formal comparisons of stochastic optimization problems with their fluid models, have been an interesting problem, as considered on various settings in Altman et al (2001); Chen and Mandelbaum (1991, 1994); Chen (1996); Dai (1995); Mandelbaum and Pats (1995); Mandelbaum et al (1998); Maglaras (2000); Gajrat et al (1997); Gajrat and Hordijk (2000); Gajrat et al (2003); Bäuerle (2000, 2002); Ethier and Kurtz (1986); Pang and Day (2007); Piunovskiy (2009a,b); Robert (2003) (again this list is far not being exhaustive). However, out of those works, Chen and Mandelbaum (1991, 1994); Chen (1996); Dai (1995); Mandelbaum and Pats (1995); Mandelbaum et al (1998) as well as (Ethier and Kurtz 1986, Chap.11) and (Robert 2003, Chap.9) only compare the trajectories, which do not directly reveal the difference between performance functionals, of particular interest in the context of optimizations, Gajrat et al (1997); Gajrat and Hordijk (2000); Gajrat et al (2003) only consider stochastic optimization problems in discrete time, Altman et al (2001); Bäuerle (2000, 2002); Maglaras (2000); Pang and Day (2007) do not provide the accuracy of fluid approximations in closed-form, and Piunovskiy (2009a,b) are, to our best knowledge, the few works in the current literature, which do provide the accuracy of fluid approximations for models like those described in Preliminary examples (a, b), based on however, somewhat restrictive conditions, requiring for example, boundedness and some version of continuity of the transition rates. In comparison, the exposition of fluid approximations in this thesis provides their accuracy, and more interestingly, notices the accuracy can change qualitatively with the values of parameters of the underlying system. Thirdly, we address possible answers to Problem set (c). The problem of buffer sizing has been studied intensively by many authors, see for example, Appenzeller et al (2004); Avrachenkov et al (2002, 2005, 2010); Dhamdhere et al (2005); Dhamdhere and Dovrolis (2006); Enachescu et al (2005); Ganjali and McKeown (2006); Lakshmikantha et al (2008); Morris (1997, 2000); Prasad et al (2007); Stanojević et al (2006); Gorinsky et al (2005); Villamizar and Song (1994); Vu-Brugier et al (2007); Raina et al (2005); Wischik and McKeown (2005); Zhang and Loguinov (2008). However, most of them focus on the network, where flows are under the AIMD (Additive Increase Multiplicative Decrease) congestion control algorithm (see Rothblum and Shorten (2007)), with the exceptional cases of Stanojević et al (2006); Zhang and Loguinov (2008), which advocate sizing buffer adaptively, and hence no formulae for the buffer size are provided. By the way, the AIMD congestion control algorithm is the cornerstone of the current TCP New Reno version. On the other hand, it has been noticed that the current TCP New Reno version is not able to utilize efficiently high speed links (see Floyd (2003)). Consequently, several more aggressive alternatives have been proposed with one example being STCP, which relies on the MIMD congestion control algorithm. In response to that, in this thesis, rather than focusing on the AIMD congestion control algorithm, we tackle the buffer sizing problem analytically for the case of the MIMD congestion control algorithm.

In the above we have revealed that Problem sets (a, b, c) are related to each other. On the other hand, to some extent, they are also independent: for example, one has no need to have in mind CTMDPs, when investigating the buffer sizing problem. For this reason, the rest of this thesis is organized as follows: the main body of this thesis is divided into three self-closed parts (in the sense of one part not directly relying on the results derived in the others), where the first part consisting of Chapter 2 focuses on Problem set (a), the second part consisting of Chapter 3, Chapter 4 and Chapter 5 is concentrated on Problem set (b), and the third part comprised of Chapter 6 responds to Problem set (c). The greater details are as follows:

Chapter 2 tackles constrained CTMDPs with the state space and action space being both Borel. The optimality criterion to be minimized is the expected discounted cost, while several constraints of the same type are imposed. The transition rates and cost rates can be unbounded (from above and from below). For the special case of no constraint, following the dynamic programming approach the existence of deterministic stationary optimal policies is shown, where the optimality is out of the class of history-dependent policies. As for constrained CTMDPs, by adopting the convex analytic approach we introduce and study occupation measures, establish the equivalent linear program formulations for the concerned optimization problems, and finally show the existence of stationary constrained-optimal policies. The power of the derived results is illustrated with an example, which, to our best knowledge, cannot be covered by the entire current literature on CTMDPs.

In Chapter 3, we consider the fluid model of a controlled Birth-and-Death process with an absorbing state, where the performance criterion is the expected total cost up to the absorption (recall Preliminary example (a)). After noticing by means of an example that the standard fluid model can fail to provide reasonable approximations (even in the sense of fluid scaling), we propose a refined fluid model for the concerned problem. For this refined fluid model we provide its accuracy, which is then verified by illustrative examples.

Results obtained in Chapter 3 also have implications on stochastic problems with a long run average expected cost. One example of such problems is the EOQ model, where, as we recall, one aims at obtaining the EOQ minimizing TCU. In Chapter 4, we are concerned with the fluid approximations to the EOQ model. Based on the results for the Birth-and-Death process with an absorbing state like those derived in Chapter 3, we provide the efficiency of a natural translation mechanism for the order quantity (particularly EOQ) in the fluid model, which is then verified by examples. As a by-product, the EPQ model is also studied.

Chapter 5 considers fluid approximations to a general bandwidth-sharing network (with an arbitrary number of flows and resources), where the concerned stochastic optimization problem is taken as a finite CTMDP (recall Preliminary example (c)). Here special attention is paid on the so called tracking policy (see Bäuerle (2000)) translating the fluid optimal policy back into one for the stochastic model, which is piecewise-constant in time (in other words, not stationary). Particularly we study the efficiency of the tracking policy. The most interesting observation is that the efficiency can change qualitatively, depending on the values of parameters.

In Chapter 6, by formulating and investigating a fluid model we study the interaction between the MIMD congestion control algorithm and a bottleneck router with Drop Tail buffer. We study conditions under which the system trajectories converge to limiting cycles with a single jump. Following that, we consider the problem of the optimal buffer sizing in the framework of multi-criteria optimization in which one accounts for the average throughput (to be maximized) and the average delay in the queue (to be minimized). As case studies, we consider the Slow Start phase of TCP New Reno and Scalable TCP for high speed networks.

Eventually we finish this thesis with conclusions and appendices.

As for references, this thesis largely comes from the following collaborative results:

- "Convergence of trajectories and optimal buffer sizing for MIMD congestion control" (with K. Avrachenkov, U. Ayesta and A. Piunovskiy, co-authors). Computer Communications, Vol 33, Issue 2, p.149-159, 2010.

- "Fluid model of an Internet router under the MIMD control scheme" (with U. Ayesta and A. Piunovskiy, co-authors). Chapter 11 in Telecommunications modeling, policy, and technology, p.239-251. editors: S. Raghavan, B. Golden and E. Wasil, Springer, NY, 2008.
- "Accuracy of fluid approximations to controlled Birth-and-Death processes: absorbing case" (with A. Piunovskiy, co-author). Submitted to Mathematical Methods of Operations Research for publication.
- "Asymptotic fluid optimality and efficiency of tracking policy for bandwidth-sharing networks" (with K. Avrachenkov and A. Piunovskiy, co-authors). Submitted to Journal of Applied Probability for publication.
- "On the fluid approximations of a class of general inventory level-dependent EOQ and EPQ model" (with A. Piunovskiy, co-author). Submitted to Central European Journal of Operations Research for publication.
- "Constrained discounted continuous-time Markov decision processes with unbounded transition and cost rates: the case of Borel state space" (with A. Piunovskiy, co-author). In preparation.

Finally, let us finish this chapter with a summary of denotations and abbreviations. The following denotations are frequently used throughout this thesis:

- $a.e.$: almost everywhere,
- $a.s.$: almost surely,
- $s.t.$: subject to,
- resp.: respectively,
- I : the indicator function,
- $\delta_x(\cdot)$: the Dirac measure concentrated at x,
- $\mathscr{B}(X)$: the Borel σ-algebra of the set X,
- $\sigma(X)$: the σ-algebra generated by a set or a random variable X,
- $\mathscr{F}_1 \bigvee \mathscr{F}_2$: the smallest σ-algebra containing the two σ-algebras $\mathscr{F}_1$ and $\mathscr{F}_2$,
- X^c : the complement of the set X,
- $|\mu|$: the total variation of the signed measure μ,
- $u(n) = O(\frac{1}{n})$: $\lim_{n\to\infty} \frac{u(n)}{1/n} = 1$,

- $\mathbb{R}_+$: $(0,\infty)$,
- $\mathbb{R}_+^0$: $[0,\infty)$,

The abbreviations of the main terminologies used in this thesis are collected as follows, too:

- AFO: Asymptotically Fluid Optimal,
- AO: Asymptotically Optimal,
- AIMD: Additive Increase Multiplicative Decrease,
- BDP: Bandwidth Delay Product,
- BS: Bandwidth-Sharing,
- CTMDP: Continuous-Time Markov Decision Processes,
- DLP: Dual Linear Program,
- DTMDP: Discrete-Time Markov Decision Processes,
- EIBL: Economic Inventory Backup Level,
- EOQ: Economic Order Quantity,
- EPQ: Economic Production Quantity,
- LHS: Left Hand Side,
- MIMD: Multiplicative Increase Multiplicative Decrease,
- PLP: Primal Linear Program,
- RHS: Right Hand Side,
- RTT: Round Trip Time,
- STCP: Scalable TCP,
- TCP/IP: Transmission Control Protocol/ Internet Protocol,
- TCU: Total Cost Per Unit Time.

Chapter 2

Constrained discounted CTMDP

2.1 Introduction

Roughly speaking, a CTMDP models a jump process, where one can control its transition rates, if necessary, at any time moment, so as to optimize a performance functional regarding a given criterion. Here the transition rates depend only on the current state of the process and the action being chosen. If apart from the key performance functional, one must ensure the acceptable performance of the system regarding some other performance criteria, then the underlying CTMDP is called a constrained one. If the performance functional(s) are discounted, the regarding (constrained) CTMDP is called a discounted one. Here we emphasize that the formal definitions of all the terminologies of appearance in this section will be given shortly in Section 2.2. From the practical point of view, the decision maker can rely on all the history about the system. Also, a decision can be potentially chosen randomly. Taking into account these two features, the standard and rigorous construction of CTMDPs was known in Kitaev (1986). An alternative construction can be found in (Guo and Hernández-Lerma 2009, Chap.2), which however, only allows one to consider at most the class of Markov policies. Here Markov policies mean that decisions are made possibly randomly, but only based on the current state and time of the system. By the way, in what follows, we often use the two terms of "CTMDP" and "CTMDP optimization problem" interchangeably.

Regarding CTMDPs, one question for study lies in whether it is necessary to make decisions randomly and based on the history, and particularly, whether it is sufficient to make decisions deterministically and based only on the current state of the system. A decision-making rule (policy) of this type is often called deterministic stationary. Another question is whether the concerned CTMDP is solvable, or in other words, whether it admits a policy, under which the system achieves the optimal performance. Both questions are important regarding the decision maker: a deterministic stationary optimal policy is clearly much "cheaper" and "easier", because one has no need to memorize the history, or to switch controls between two consecutive jumps of the controlled process.

Revolving about these two questions, initial studies of discounted CTMDPs are mainly concentrated on the case of uniformly bounded transition rates. On such occasions, one may indeed apply the well known uniformization technique (Puterman (1994)) to reduce the continuous time problem to an equivalent discrete time problem. Then, relatively well-developed theories on DTMDPs (Discrete-Time Markov Decision Processes) (see for example, Altman (1999); Piunovskiy (1997)) can be applied. For example, Feinberg (2004); Piunovskiy (1998, 2005) studied the solvability and the form of optimal policies for constrained discounted CTMDPs with a Borel and countable state space, respectively. Indeed, to our best knowledge, Piunovskiy (1998) is one of the first papers in constrained CTMDPs. The more challenging case is when the transition rates may be unbounded. For instance, based on the dynamic programming approach, Guo and Dai (2002); Guo et al (2006); Guo (2007); Guo and Hernández-Lerma (2003c); Guo and Zhu (2002); Yan et al (2008), with Guo and Dai (2002); Guo et al (2006); Guo and Hernández-Lerma (2003c); Guo and Zhu (2002) on countable state space and Guo (2007); Yan et al (2008) on Polish state space, showed the existence of deterministic stationary optimal policies, but the optimality is only out of a special class of Markov policies. Note, Guo and Dai (2002); Guo et al (2006); Guo (2007); Guo and Hernández-Lerma (2003c); Guo and Zhu (2002); Yan et al (2008) are all about unconstrained discounted CTMDPs only. For the constrained discounted CTMDPs, the dynamic programming approach was also applied in Guo and Hernández-Lerma (2003a) to study the solvability as well as the form of optimal policies. However, over there the authors only considered one constraint, and again the optimality is out of the special class of Markov policies. Indeed, when the optimality is out of the class of history-dependent policies, the study of CTMDPs with possibly unbounded transition rates had been long known as a challenging one, and the most recent survey of the theory of CTMDPs Guo et al (2006) still mentioned it as an open problem. To our best knowledge, the first successful treatment of this open problem is the recent work Guo and Piunovskiy (2010), where the authors studied constrained discounted CTMDPs for the case of countable state space, based on the combination of the dynamic programming approach as well as the convex analytic approach.

In a nutshell, to our best knowledge, this chapter is the second treatment of the open problem mentioned in the previous paragraph. In more details, this chapter is about constrained discounted CTMDPs with Borel state space and action space, where the transition rates and cost rates can be unbounded. Our main contributions are threefolded: firstly, for unconstrained CTMDPs, which can be regarded as a special case of constrained CTMDPs, we derive the corresponding Bellman equation, prove the existence of deterministic stationary optimal policies; secondly, for the general case of constrained CTMDPs, we study the space of occupation measures, and establish equivalent linear program formulations to our optimization problems, for which we prove the absence of the duality gap between the PLP (Primal Linear Program) and its DLP (Dual Linear Program); and thirdly, we prove the solvability of the constrained CTMDPs, and show the existence of a (possibly randomized) stationary optimal policy. By the way, it is known that in general the class of deterministic stationary policies are not sufficient for constrained CTMDPs, see (Piunovskiy 1998, Sec.7). Here we emphasize that our optimality is out of the class of history-dependent policies, in which sense, even for the unconstrained CTMDPs, when the state space is Borel and the transition rates are unbounded, the sufficiency of the class of deterministic stationary policies has not been shown in the current literature. In relation to the most closely related works, for unconstrained CTMDPs, the current work complements Guo (2007); and for constrained CTMDPs, it extends Guo and Piunovskiy (2010) to the case of Borel state space.

The rest of this chapter is organized as follows: in Section 2.2, we remind the construction of the concerned constrained discounted CTMDPs, define the main terminologies and present some preliminary results. In Section 2.3 we focus on the unconstrained CTMDPs, and develop the dynamic programming approach to derive the existence of deterministic stationary optimal policies, while in Section 2.4 for the constrained case, we develop the convex analytic approach to derive the existence of stationary constrained-optimal policies. Section 2.5 provides an example to illustrate a situation, where all our theorems are applicable, but those in all the previous ones in the theory of CTMDPs are not. We finish this chapter with conclusions in Section 2.6. The proofs of the main statements are collected in Section 2.7 at the very end.

2.2 Preliminaries

2.2.1 Kitaev's construction of the controlled process

We would like to make it clear that the materials presented in this subsection are from Kitaev (1986); Kitaev and Rykov (1995); Piunovskiy (1998), but they are necessary to introduce all the notations.

The primitives of the concerned CTMDPs are the following elements:

- state space: $(S,\mathscr{B}(S))$,
- action space: $(A,\mathscr{B}(A))$,
- admissible action space $A(x)\in\mathscr{B}(A)$ and the space of admissible action-state pairs $K\triangleq\{(x,a)\in S\times A : a\in A(x)\}\in\mathscr{B}(S\times A)$, assumed to contain the graph

of a measurable function ϕ from S to A such that $\forall\, x \in S,\ \phi(x) \in A(x)$,

- transition rate: $q(dy|x,a)$, a signed kernel on $\mathscr{B}(S)$ given $(x,a) \in K$, taking nonnegative values on $\Gamma_S \setminus \{x\}$ with $\Gamma_S \in \mathscr{B}(S)$, being conservative in the sense of $q(S|x,a) = 0$ and stable in that $\bar{q}_x = \sup_{a \in A(x)} q_x(a) < \infty$, where $q_x(a) \triangleq -q(\{x\}|x,a)$,
- key cost rate: $c_0(x,a)$ measurable in $(x,a) \in K$,
- sources of constraints: $(c_i(x,a), d_i)_{i=1,\dots,N}$, where for any $i = 1,\dots,N$, $c_i(x,a)$ is measurable in $(x,a) \in K$, $d_i \in \mathbb{R}$, and N is the number of constraints,
- discount factor: $\alpha > 0$,
- initial distribution: $\gamma(\cdot)$, a probability measure on $(S, \mathscr{B}(S))$.

Incidentally speaking, we remind that a singleton $\{x\} \subseteq S$ is measurable; and $q_x(a)$ is measurable on K, see (Bertsekas and Shreve 1978, Prop 7.29). In what follows, for the sake of formality, if needed, for any $\Gamma_S \in \mathscr{B}(S)$, we may consider $q(\Gamma_S|x,a)$ as its measurable extension on $S \times A$, where $q(\Gamma_S|x,a) = 0$ on $(S \times A) \setminus K$, and similar assertions are applicable to other functions such as c_n, and so on. This is just the convention (Hernández-Lerma and Lasserre 1996, Chap.6).

Given the above primitives, let us now remind the readers of the construction of the underlying stochastic basis $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, P^{\pi}_{\gamma})$ and the controlled process $\{\xi_t, t \geq 0\}$ thereon, as given in Kitaev (1986) and Piunovskiy (1998), and this will be done in steps.

Step 1: measurable space $(\Omega, \mathscr{F})$. Starting with the measurable space $(\Omega^0, \mathscr{F}^0) \triangleq ((S \times \mathbb{R}_+)^{\infty}, \mathscr{B}((S \times \mathbb{R}_+)^{\infty}))$, let us adjoin all the sequences of the form

$$(x_0,\ \theta_1,\ x_1,\ \dots,\ \theta_{m-1},\ x_{m-1},\ \infty,\ x_\infty\ ,\infty,\ x_\infty,\ \dots)$$

to Ω^0, where $x_0 \in \mathscr{B}(S)$, $x_\infty \notin S$ is an isolated point, $m \geq 1$ is some integer (we omit θ_0 in case of $m = 1$), $\theta_l \in \mathbb{R}_+$ and $x_l \neq x_\infty$ for all nonnegative integers $l \leq m-1$. After the corresponding modification of the σ-algebra $\mathscr{F}^0$, we obtain the basic measurable space $(\Omega, \mathscr{F})$.

Step 2: stochastic process $\{\xi_t, t \geq 0\}$ and history $\{\mathscr{F}_t\}_{t\geq 0}$. Putting $T_0 \triangleq 0$, $T_m \triangleq \theta_1 + \theta_2 + \dots + \theta_m$, $T_\infty \triangleq \lim_{m\to\infty} T_m$, we can define the process of interest:

$$\xi_t(\omega) \triangleq \sum_{m\geq 0} I\{T_m \leq t < T_{m+1}\} x_m + I\{T_\infty \leq t\} x_\infty$$

together with the history it is adapted to:

$$\mathscr{F}_t \triangleq \sigma(\{T_m \leq s, x_m \in \Gamma_S\} : \Gamma_S \in \mathscr{B}(S), s \leq t, m \geq 0).$$

In what follows, $\omega = \{x_0, \theta_1, x_1, \ldots\}$ is often omitted, and $h_m(\omega) = (x_0, \ldots, \theta_m, x_m)$ is referred to as an m-component history. Here θ_m (resp. T_m, x_m) can be understood as the inter-jump intervals or sojourn times (resp. the jump moments, the state of the process on the interval $[T_m, T_{m+1})$). We shall not intend to consider the process after T_∞ : the isolated point x_∞ will be regarded as absorbing.

Step 3: policy π. Having adjoint the isolated point a_∞ to A, we thus define $A_\infty \triangleq A\bigcup\{a_\infty\}$, and put $A(x_\infty) \triangleq \{a_\infty\}$. Similarly, S_∞ can be understood. Denoting $\mathscr{F}_{s-} \triangleq \bigvee_{t<s}\mathscr{F}_t$, the predictable (with respect to $\{\mathscr{F}_t\}_{t\geq 0}$) σ-algebra $\mathscr{P}$ on $\Omega \times \mathbb{R}^0_+$ is given by

$$\mathscr{P} \triangleq \sigma(\Gamma_S \times \{0\}\ (\Gamma \in \mathscr{F}_0), \Gamma \times (s,\infty)\ (\Gamma \in \mathscr{F}_{s-})).$$

See (Kitaev and Rykov 1995, Chap.4.4) for more details. Now the following series of definitions are in position:

- History-dependent policy: $\pi(\cdot|\omega,t)$, a $\mathscr{P}$-measurable transition probability function on $(A_\infty, \mathscr{B}(A_\infty))$, concentrated on $A(\xi_{t-})$.
- Markov policy: $\pi(\cdot|\xi_{t-},t) = \pi(\cdot|\omega,t)$. Here concerning the LHS, $\pi(\cdot|x,t)$ is $\mathscr{B}(S_\infty \times \mathbb{R}^0_+)$-measurable.
- Stationary policy: $\pi(\cdot|\xi_{t-}(\omega)) = \pi(\cdot|\omega,t)$. Here concerning the LHS, $\pi(\cdot|x)$ is $\mathscr{B}(S_\infty)$-measurable.
- Deterministic stationary policy: $\phi(\xi_{t-}(\omega)) = \phi(\omega,t)$, where $\{\phi(\xi_{t-}(\omega), t\geq 0)\}$ is a $\mathscr{P}$-measurable A_∞-valued process such that $\pi(\Gamma^A|\omega,t) = I\{\Gamma^A \ni \phi(\omega,t)\}$ for all $\Gamma^A \in \mathscr{B}(A_\infty)$. Here $\phi(\cdot)$ is $\mathscr{B}(S_\infty)$-measurable, and we use ϕ rather than π to signify the "deterministicity".

Note, in the current setup, a deterministic stationary policy always exists, see (Hernández-Lerma and Lasserre 1996, p.14).

Step 4: (γ,π-dependent) probability measure P^π_γ on $(\Omega, \mathscr{F})$. Under any fixed policy π, let us define

$$\nu^\pi(\omega, \Gamma_S \times dt) \triangleq \Lambda(\Gamma_S|\omega,t)dt \triangleq \left[\int_A \pi(da|\omega,t)q(\Gamma_S \setminus \{\xi_{t-}\}|\xi_{t-},a)\right] dt, \qquad (2.1)$$

where $\Gamma_S \in \mathscr{B}(S)$, and the obvious dependence of Λ on π has been omitted. This random measure is predictable, see Kitaev (1986); Kitaev and Rykov (1995); Piunovskiy (1998). In fact, the "jump intensity" Λ has the following form (Jacod 1975, Eqn.(6)):

$$\begin{aligned}&\Lambda(dy|\omega,t)\\ =\ &\sum_{m\geq 0} I\{T_m < t \leq T_{m+1}\}\Lambda^m(dy|x_0,\ldots,x_m,t-T_m) + I\{t=0\}\Lambda^0(dy|x_0), \quad (2.2)\end{aligned}$$

where for fixed $\Gamma_S \in \mathscr{B}(S)$, $\Lambda^m(\Gamma_S|x_0,\theta_1,\dots,x_m,u)$ are some non-negative, non-random measurable functions. See (Jacod 1975, Lem.3.3) for more details. Then comparing (2.1) with (2.2), we have the explicit formula for Λ^m :

$$\Lambda^m(dy|x_0,\dots,x_m,u) = \int_A \pi(da|x_0,\dots,x_m,u+T_m)q(dy\setminus\{x_m\}|x_m,a). \qquad (2.3)$$

Let $\hat{H}_0 \triangleq S$ and $\hat{H}_m \triangleq S\times((0,\infty]\times S_\infty)^m, m=1,2,\dots$. Regulate that on $\hat{H}_0$, P_γ^π is given to be γ. Suppose now, P_γ^π on $\hat{H}_m$ for $1\le m\le k$ is constructed. Here without leading to confusion, we have abused the denotation of P_γ^π by considering its marginal. It is needed now to construct P_γ^π on $\hat{H}_{k+1}$. But this can be done via

$$\begin{aligned} & P_\gamma^\pi(\Gamma^{\hat{H}_k}\times(du\times dy)) \\ \triangleq\; & \int_{\Gamma^{\hat{H}_k}} P_\gamma^\pi(dh_k)I\{\theta_k<\infty\}\Lambda^k(dy|h_k,u)\times e^{-\int_0^u \Lambda^k(S|h_k,v)dv}du, \\ & P_\gamma^\pi(\Gamma^{\hat{H}_k}\times(\infty,x_\infty)) \\ \triangleq\; & \int_{\Gamma^{\hat{H}_k}} P_\gamma^\pi(dh_k)\left\{I\{\theta_k=\infty\} + I\{\theta_k<\infty\}e^{-\int_0^\infty \Lambda^k(S|h_k,v)dv}\right\}, \end{aligned} \qquad (2.4)$$

where $\Gamma^{\hat{H}_k}\in\mathscr{B}(\hat{H}_k)$. Now it only remains to apply the induction and Ionescu-Tulcea's theorem (Bertsekas and Shreve 1978, p.140-141, Prop.7.28) to induce that P_γ^π is the unique probability measure on $(\Omega,\mathscr{F})$ such that its projection (marginal) onto $\hat{H}_m$ satisfies (2.4), $m=0,1,\dots$. This thus gives rise to $(\Omega,\mathscr{F},\{\mathscr{F}_t\}_{t\ge0}\},P_\gamma^\pi)$, always assumed to be complete, which thus completes Kitaev's construction.

In fact, according to Kitaev (1986), if we define the random measure

$$\mu(\omega,dt,dy) \triangleq \sum_{m\ge1} I\{T_m<\infty\}I\{x_m\in dy\}I\{T_m\in dt\}, \qquad (2.5)$$

then under any fixed policy π and for any given initial distribution γ, the above defined P_γ^π on $(\Omega,\mathscr{F})$ is such that its projection on the 0-component history is γ, and ν^π defined in (2.1) is a dual predictable projection of μ defined in (2.5). See Kitaev and Rykov (1995) for more details.

Below, when $\gamma(\cdot)$ takes the form of the Dirac measure concentrated on $x\in S$, we shall use the "degenerated" denotation P_x^γ. Expectations with respect to P_γ^π and P_x^π are denoted as E_γ^π and E_x^π, respectively.

Remark 2.1 *Equality (2.3) holds P_γ^π-a.s., as well as all the subsequent equalities and inequalities involving ω.*

2.2.2 Properties of the controlled process

Lemma 2.1 *Let constants $\rho \neq 0$, $b \geq 0$, measurable function $w(x) \geq 0$ on S and a signed kernel $f(dy|x,t)$ on $\mathscr{B}(S)$ given $(x,t) \in S \times \mathbb{R}^0_+$ be such that*
(a) $f(\Gamma_S|x,t) \geq 0 \; \forall \; \Gamma_S \in \mathscr{B}(S)$ such that $\Gamma_S \bigcap \{x\} = \emptyset$, $F_x(t) \triangleq f(S \setminus \{x\}|x,t) < \infty$, and $f(S|x,t) = 0$;
(b) $\int_S f(dy|x,t)w(y) \leq \rho w(x) + b$.

Define the non-negative function

$$h(s,x,t) \triangleq e^{\rho(t-s)}w(x) + \frac{b}{\rho}(e^{\rho(t-s)} - 1), \tag{2.6}$$

where $s \leq t$, $s,t \in \mathbb{R}^0_+$ and $x \in S$. Then the following inequality holds:

$$h(s,x,t) \geq \int_s^t \int_{S\setminus\{x\}} \exp(-\int_s^u F_x(v)dv) f(dy|x,u)h(u,y,t)du + \exp(-\int_s^t F_x(v)dv)w(x).$$

Conditions 2.1 *There exist a measurable (weight) function $w(x) \geq 1$ on S and constants $\rho \geq 0$ and $b \geq 0$ such that*
(a) There exist an increasing system of measurable $S_n \subseteq S$ such that $\bigcup_{n=0}^{\infty} S_n = S$, $\sup_{x \in S_n} w(x) < \infty$ as well as $\lim_{n\to\infty} \inf_{x \in S\setminus S_n} w(x) = \infty$.
(b) $\int_S q(dy|x,a)w(y) \leq \rho w(x) + b$ for all $x \in S$, $a \in A(x)$;
(c) For any $l \geq 0$, $\sup_{x \in S_l} \bar{q}_x < \infty$, where S_l has been defined in part (a), and $\bar{q}_x \triangleq \sup_{a \in A(x)} q_x(a)$.

Function $\bar{q}_x$ is measurable under Conditions 2.4 for example, see below. Part (b) of Conditions 2.1 was also assumed in Guo (2007). Moreover, instead of part (c) of Conditions 2.1, to prove the regularity (see Theorem 2.1 below) and study the dynamic programming approach (see Section 2.3 below), the following stronger version was assumed in Guo (2007):

Conditions 2.2 *There exists a constant $L > 0$ such that $0 \leq \bar{q}_x < Lw(x)$, with $x \in S$ being arbitrarily fixed.*

Lemma 2.2 *Suppose Conditions 2.1 (b) are satisfied, thus fixing $w(x) \geq 1$, $b \geq 0$ and $\rho \geq 0$. Then under any π, for each $x \in S$ and $m = 0, 1, 2, \ldots$,*

$$E_x^\pi[w(\xi_t)I\{t < T_{m+1}\}] \leq (e^{\rho t}w(x) + \frac{b}{\rho}(e^{\rho t} - 1))I\{\rho > 0\} + (w(x) + bt)I\{\rho = 0\}.$$

Theorem 2.1 *Let Conditions 2.1 be satisfied. Then for any π, $x \in S$ and $t \in \mathbb{R}^0_+$, the following assertions hold:*
(a) $P_x^\pi(T_\infty = \infty) = 1$.
(b) $E_x^\pi[w(\xi_t)] \leq I\{\rho > 0\}h(0,x,t) + I\{\rho = 0\}(w(x) + bt)$, where h is defined in (2.6).
(c) $P_x^\pi(\xi_t \in S) = 1$.

Evidently, part (a) of Theorem 2.1 is equivalent to that for any π, $x \in S$ and $t \in \mathbb{R}^0_+$, $P^\pi_\gamma(T_\infty = \infty) = 1$; whereas parts (b) and (c) were proved in Guo (2007) for the class of Markov policies.

Remark 2.2 *From now on, without loss of generality, we shall assume $\rho > 0$, where ρ is defined in Conditions 2.1, because the case of $\rho = 0$ can always be considered by passing to the limit as $\hat{\rho} \to 0$, with $\hat{\rho} > 0$.*

Conditions 2.3 *There exist constants $M \geq 0$ and c such that*
(a) $\int_S \gamma(dy)w(y) < \infty$, where γ is the given initial distribution.
(b) $\alpha > \rho$, where α is the discount factor, and ρ is as in Conditions 2.1.
(c) $|c_n(x,a)| \leq Mw(x) + c$ for $(x,a) \in K$ and $n = 0,\ldots,N$.

Note, parts (b) and (c) of Conditions 2.3 were also assumed in Guo (2007), where the author restricted himself to γ being δ_x, and thus did not require part (a).

Now for $n = 0,1,\ldots,N$, the following two quantities are well defined in Corollary 2.1 below:

$$V_n(x,\pi) \triangleq E^\pi_x\left[\int_0^\infty e^{-\alpha t}\int_A c_n(\xi_{t-},a)\pi(da|\omega,t)dt\right] \tag{2.7}$$

and

$$V_n(\pi) \triangleq E^\pi_\gamma\left[\int_0^\infty e^{-\alpha t}\int_A c_n(\xi_{t-},a)\pi(da|\omega,t)dt\right] = \int_S V_n(x,\pi)\gamma(dx), \tag{2.8}$$

For notational convenience, below we shall often replace ξ_{t-} with ξ_t in formulae like (2.7) and (2.8) without generating any confusion.

Corollary 2.1 *Suppose Conditions 2.1 and Conditions 2.3 are satisfied. Then*

$$|V_n(x,\pi)| \leq \frac{M(\alpha w(x) + b)}{\alpha(\alpha-\rho)} + \frac{c}{\alpha} < \infty$$

and

$$|V_n(\pi)| \leq \frac{M(\alpha\int_S \gamma(dy)w(y) + b)}{\alpha(\alpha-\rho)} + \frac{c}{\alpha} < \infty.$$

Before stating the next result, let us impose conditions for the measurability of $\bar{q}_x$.

Conditions 2.4 *(a) $A(x)$ is compact for any $x \in S$.*
(b) $q_x(a)$ is upper semicontinuous on $A(x)$ for any fixed $x \in S$.

Note, $\bar{q}_x$ has appeared in the proof of previous statements, where however, it is not necessarily measurable. Then according to (Hernández-Lerma and Lasserre 1996, D.5 Prop.) (see also (Bertsekas and Shreve 1978, Prop.7.33)), under Conditions 2.4, $\bar{q}_x$ is measurable on S.

Remark 2.3 *To ensure the measurability of $\bar{q}_x$, alternatively to Conditions 2.4, the following set of conditions can be imposed:*

Conditions 2.5 *(a) K is an open subset of $S \times A$.*
(b) $q_x(a)$ is lower semicontinuous on K.

According to (Bertsekas and Shreve 1978, Prop.7.34), under Conditions 2.5, $\bar{q}_x$ is measurable on S. Consequently, results requiring Conditions 2.4 hold also, if we instead impose Conditions 2.5.

Remark 2.4 *If Conditions 2.4 (a) (resp. Conditions 2.5 (a)) hold, then according to (Hernández-Lerma and Lasserre 1996, D5 Prop.) (resp. (Bertsekas and Shreve 1978, Prop.7.34)), there is a deterministic stationary policy.*

Theorem 2.2 *Let Conditions 2.1 and Conditions 2.4 (or Conditions 2.5) be satisfied. Then for any π, $x \in S$ and $t \in \mathbb{R}^0_+$, the following analogue to the Kolmogorov's forward equation (in the integral form) holds: $\forall\, \Gamma \in \mathscr{B}(S)$ such that $\exists\, l : \Gamma \subseteq S_l$, with S_l defined in Conditions 2.1,*

$$\begin{aligned} P_x^\pi(\xi_t \in \Gamma) &= I\{x \in \Gamma\} + E_x^\pi\left[\int_0^t \int_A \pi(da|\omega,u) q(\Gamma \setminus \{\xi_{u-}\}|\xi_{u-},a)du\right] \\ &\quad - E_x^\pi\left[\int_0^t \int_A \pi(da|\omega,u) q_{\xi_{u-}}(a) I\{\xi_{u-} \in \Gamma\}du\right], \end{aligned} \tag{2.9}$$

where both the expectations in the above expression are finite.

Note, the process under an arbitrary policy π can be not Markov.

Remark 2.5 *If we replace part (c) of Conditions 2.1 by Conditions 2.2, then in Theorem 2.2, we have that for any $\Gamma \in \mathscr{B}(S)$*

$$\begin{aligned} P_x^\pi(\xi_t \in \Gamma) &= I\{x \in \Gamma\} + E_x^\pi\left[\int_0^t \int_A \pi(da|\omega,u) q(\Gamma \setminus \{\xi_{u-}\}|\xi_{u-},a)du\right] \\ &\quad - E_x^\pi\left[\int_0^t \int_A \pi(da|\omega,u) q_{\xi_{u-}}(a) I\{\xi_{u-} \in \Gamma\}du\right], \end{aligned} \tag{2.10}$$

where both the expectations in the above expression are finite. Indeed, one only needs replace the argument for (2.40) in the proof of Theorem 2.2 by the following:

$$\begin{aligned} E_x^\pi[\tilde{\mu}((0,t],\Gamma)] &= E_x^\pi[\tilde{\nu}((0,t],\Gamma)] = E_x^\pi\left[\int_0^t \int_A \pi(da|\omega,u) q_{\xi_{u-}}(a) I\{\xi_{u-} \in \Gamma\}du\right] \\ &\le E_x^\pi\left[\int_0^t \bar{q}_{\xi_{u-}} I\{\xi_{u-} \in \Gamma\}du\right] \le E_x^\pi\left[\int_0^t Lw(\xi_{u-}) I\{\xi_{u-} \in \Gamma\}du\right] \\ &\le \int_0^t E_x^\pi[Lw(\xi_u)]\,du \le \int_0^t Lh(0,x,u)du < \infty, \end{aligned}$$

where the second inequality follows from part (c) of Conditions 2.1, the fourth inequality follows from part (b) of Theorem 2.1, and h is defined in (2.6) (see also Remark 2.2).

This analogue to the Kolmogorov's forward equation has been proved in (Kitaev 1986, Lem.4) for the case of uniformly bounded $\bar{q}_x$.

Conditions 2.6 *There exist a measurable function $w'(x) \geq 1$ on S and non-negative constants L', ρ' and b' such that the following assertions hold:*
(a) $(\bar{q}_x + 1)w'(x) \leq L'w(x)$, where w comes from Conditions 2.1;
(b) $\int_S q(dy|x,a)w'(y) \leq \rho' w'(x) + b'$ for all $x \in S$, $a \in A(x)$;
(c) $\alpha > \rho'$;
(d) $|c_n(x,a)| \leq Mw'(x) + c$ for all $(x,a) \in K$ and $n = 0, 1, \ldots, N$.

Conditions 2.6 are essentially equivalent to (Guo 2007, Assump.b and c(4)).

Definition 2.1 *Functions w and w' coming from Conditions 2.1 and Conditions 2.6 are called weight functions, and any measurable function u on S such that $\sup_{x\in S} \frac{|u(x)|}{w(x)} < \infty$ (resp. $\sup_{x\in S} \frac{|u(x)|}{w'(x)} < \infty$) is called to have a bounded w- (resp. w'-) weighted norm, with the norm $||u||_w \triangleq \sup_{x\in S} \frac{|u(x)|}{w(x)}$ (resp. $||u||_{w'} \triangleq \sup_{x\in S} \frac{|u(x)|}{w'(x)}$). The collection of all functions u on S with a bounded w- (resp. w'-) weighted norm will be denoted by $\boldsymbol{B}_w(S)$ (resp. $\boldsymbol{B}_{w'}(S)$).*

Theorem 2.3 *Under Conditions 2.1, Conditions 2.4 (or Conditions 2.5) and parts (a) and (b) of Conditions 2.6, given x and π being fixed, for any function $u \in \boldsymbol{B}_{w'}(S)$, the following two versions of Dynkin's formula hold:*

$$E_x^\pi[u(\xi_t)] - u(x) = E_x^\pi\left[\int_0^t \int_S \int_A \pi(da|\omega,v)q(dy|\xi_v,a)u(y)dv\right]; \tag{2.11}$$

$$\begin{aligned} E_x^\pi[u(\xi_t)]e^{-\alpha t} - u(x) &= E_x^\pi\left[\int_0^t e^{-\alpha v}\{-\alpha u(\xi_v)\right. \\ &\quad \left. + \int_S \int_A \pi(da|\omega,v)q(dy|\xi_v,a)u(y)\Big\}dv\right]. \end{aligned} \tag{2.12}$$

2.2.3 CTMDP optimization problem statement

We are interested in the following optimization problem:

$$V_0(\pi) \to \min_\pi \tag{2.13}$$

s.t.

$$V_n(\pi) \leq d_n, \ n = 1, 2, \ldots, N.$$

In case $N = 0$ (resp. $N \in \mathbb{N}$), the CTMDP is unconstrained (resp. constrained). Let us denote $U \triangleq \{\pi : V_n(\pi) \leq d_n, n = 1, \ldots, N\}$ the set of feasible policies, assumed to be nonempty throughout this chapter, and $V_0^* \triangleq \inf_{\pi\in U} V_0(\pi)$ the constrained-optimal value.

Definition 2.2 *A policy $\pi^* \in U$ is called constrained-optimal if the infimum in (2.13) is achieved at it: $V_0(\pi^*) = V_0^*$. The CTMDP (2.13) is solvable if π^* exists.*

2.3 Dynamic programming approach for unconstrained CTMDPs: $N = 0$

In this section, we only consider the case of $N = 0$, that is, the unconstrained problem. The obtained result, of interest in its own right, is also useful in studying constrained problems in the next section.

Lemma 2.3 *Suppose Conditions 2.1, Conditions 2.3 (b) and (c), Conditions 2.4 (or Conditions 2.5), and parts (a) and (b) of Conditions 2.6 are satisfied. Then given any fixed Markov policy π and $x \in S$, the following assertions hold:*
(a) If $u \in \boldsymbol{B}_{w'}(S)$, and

$$\alpha u(x) \geq \int_A \pi(da|x,t)c_0(x,a) + \int_S \int_A \pi(da|x,t)q(dy|x,a)u(y)$$

holds for all $x \in S$ and $t \geq 0$, then $u(x) \geq V_0(x,\pi)$.
(b) If $u \in \boldsymbol{B}_{w'}(S)$, and

$$\alpha u(x) \leq \int_A \pi(da|x,t)c_0(x,a) + \int_S \int_A \pi(da|x,t)q(dy|x,a)u(y)$$

holds for all $x \in S$ and $t \geq 0$, then $u(x) \leq V_0(x,\pi)$ for all $s \in S$.

See (Guo 2007, Lem.5.3) for a proof.

Conditions 2.7 *(a) For any bounded non-negative measurable $u(y)$ on S, $u'(x,a) \triangleq \int_S u(y)q(dy|x,a)$ is lower semicontinuous in a for any fixed $x \in S$.*
(b) $\int_S w(y)q(dy|x,a)$ is continuous in $a \in A(x)$, for any fixed $x \in S$, where w comes from Conditions 2.1.
(c) $c_n(x,a)$ is lower semicontinuous in $a \in A(x)$ for any fixed $x \in S$. Here $n = 0, 1, \ldots, N$.
(d) $A(x)$ is compact for any $x \in S$.

Remark 2.6 *(a) Reasoning as in (Hernández-Lerma and Lasserre 1999, p.44), one can show that part (a) of Conditions 2.7 is equivalent to if we require for any bounded measurable (real-valued) function $u(y)$ on S, fixing some $x \in S$, $\int_S u(y)q(dy|x,a)$ is continuous in a.*
(b) Part (a) of Conditions 2.7 suffices for part (b) of Conditions 2.4.
(c) Conditions 2.7 are some sort of compactness-continuity conditions, and have been widely assumed in works about Markov decision processes on general spaces, see (Guo and Rieder 2006, Lem.3.5), as well as Guo (2007); Hernández-Lerma and Lasserre (1999); Yan et al (2008).

(d) Other than Conditions 2.7, one can impose the following set of conditions:

Conditions 2.8 *(a) For any w-bounded upper semicontinuous function u on S,* $\int_S q(dy|x,a)u(y)$ *is upper semicontinuous on K.*
(b) $c_0(x,a)$ *is upper semicontinuous on K.*
(c) $\bar{q}_x$ *is continuous on S.*
(d) K is an open subset of $S \times A$.
(e) $w(x)$ *and* $w'(x)$ *are both upper semicontinuous on S.*

Then Lemma 2.4 below holds also under Conditions 2.8, instead of Conditions 2.7, if in the statement we further require $u \in B_w(S)$ *to be upper semicontinuous. Indeed, one only needs argue similarly as in (Hernández-Lerma and Lasserre 1999, p.66-67) and apply (Bertsekas and Shreve 1978, Prop.7.34). Consequently, after Lemma 2.4, all the results requiring Conditions 2.7 in this section hold, if we instead have Conditions 2.8.*

Lemma 2.4 *Suppose Conditions 2.7 are satisfied. Then for any* $u \in \boldsymbol{B}_w(S)$, *the following function v (while it is well defined) is measurable:*

$$v(x) \triangleq \inf_{a \in A(x)} \left\{ \frac{c_0(x,a)}{\alpha + 1 + \bar{q}_x} + \frac{1 + \bar{q}_x}{\alpha + 1 + \bar{q}_x} \int_S u(y) \left(\frac{q(dy|x,a)}{1 + \bar{q}_x} + I\{x \in dy\} \right) \right\}.$$

Note that in the statement of Lemma 2.4, we can take particularly $u = w'$, where w' comes from Conditions 2.6 (a).

Assisted by Lemma 2.4, the following proposition was basically established in (Guo 2007, Thm.3.3 (b)).

Proposition 2.1 *Suppose part (b) of Conditions 2.1, parts (b) and (c) of Conditions 2.3, Conditions 2.7 (or Conditions 2.8) are satisfied. Then the following assertions hold:*
(a) The following (discounted) Bellman equation (also known as the dynamic programming equation or the optimality equation) has a solution $u^* \in \boldsymbol{B}_w(S)$:

$$\alpha u(x) = \inf_{a \in A(x)} \left\{ c_0(x,a) + \int_S q(dy|x,a)u(y) \right\}. \tag{2.14}$$

(b) The solution u^* *is the point-wise limit of the following non-increasing sequence of measurable functions* $\{u^{(n)}, n = 0, 1, \ldots\}$:

$$u^{(0)}(x) \triangleq \frac{M(\alpha w(x) + b)}{\alpha(\alpha - \rho)} + \frac{c}{\alpha},$$

$$u^{(n+1)}(x) \triangleq \inf_{a \in A(x)} \left\{ \frac{c_0(x,a)}{\alpha + 1 + \bar{q}_x} + \frac{1 + \bar{q}_x}{\alpha + 1 + \bar{q}_x} \int_S u^{(n)}(y) \left(\frac{q(dy|x,a)}{1 + \bar{q}_x} + I\{x \in dy\} \right) \right\}.$$

Remark 2.7 *Suppose Conditions 2.6 (b), (c) and (d), and Conditions 2.7 (or Conditions 2.8) are satisfied. Then the statement in Proposition 2.1 still holds, with w, ρ and b being replaced by w', ρ' and b' everywhere.*

Lemma 2.5 *Suppose Conditions 2.1, parts (a) and (b) of Conditions 2.3, Conditions 2.6 and Conditions 2.7 (or Conditions 2.8) are satisfied. Then for any π and γ, the following assertion holds:*

$$\begin{aligned} V_0(\pi) &= \int_S \gamma(dy)u^*(y) + E_\gamma^\pi \left[\int_0^\infty e^{-\alpha t} \int_A \pi(da|\omega,t) \{c_0(\xi_t,a) - \alpha u^*(\xi_t) \right. \\ &\quad \left. + \int_S q(dy|\xi_t,a)u^*(y) \Big\} dt \right], \end{aligned} \tag{2.15}$$

where $u^*(x)$ is as in Proposition 2.1 (a) and Remark 2.7. Indeed, the statement holds for any $u^* \in \mathbf{B}_{w'}(S)$.

Theorem 2.4 *Suppose Conditions 2.1, parts (a) and (b) of Conditions 2.3, Conditions 2.6 and Conditions 2.7 (or Conditions 2.8) are satisfied. Then the following assertions hold:*

(a) Let ϕ stand for any deterministic stationary policy. If $u^* \in \mathbf{B}_{w'}(S)$ solves the Bellman equation (2.14), then

$$\int_S \gamma(dy)u^*(y) = \inf_\pi V_0(\pi) = \inf_\phi V_0(\phi).$$

(b) u^* as given in Remark 2.7 (see also Proposition 2.1) is the unique solution out of $\mathbf{B}_{w'}(S)$ to the Bellman equation (2.14).

(c) u^* solves the following DLP on the space of measurable functions on S:

$$\int_S \gamma(dy)v(y) \to \max_v \tag{2.16}$$

s.t.

$$\frac{1}{\alpha}c_0(x,a) - v(x) + \frac{1}{\alpha}\int_S v(y)q(dy|x,a) \geq 0, \forall\ (x,a) \in K;$$
$$v \in \mathbf{B}_{w'}(S).$$

(d) Suppose v is feasible for the DLP (2.16). Then it solves the DLP if and only if $v(x) = u^*(x)$ $\gamma -a.s.$.

One can check that Conditions 2.1, parts (a) and (b) of Conditions 2.3, Conditions 2.6 and Conditions 2.7 cover (Guo 2007, Assump.1,2,3) (with "upper semicontinuous" being replaced by "lower semicontinuous", because here we are considering a minimization problem), validating (Guo 2007, Thm.3.3), asserting that there is a deterministic stationary policy, which is optimal for the unconstrained problem, out of a class of Markov policies, and that class is richer than the class of stationary policies. Hence, one consequence of Theorem 2.4 is the existence of a deterministic stationary optimal policy for the unconstrained problem. Note, even for the unconstrained problem, when the transition rates are possibly unbounded like in our setting, the optimality (out of the class of general history-dependent policies) and existence of deterministic stationary policies have not been established, to our best knowledge.

Remark 2.8 *In the case of countable state space, without Conditions 2.4, Conditions 2.5, Conditions 2.7 and Conditions 2.8 being satisfied, all the results derived in this section still hold.*

Let us finish this section with the following lemma to be used in the next section, which, in the case of countable state space, was established in (Guo and Hernández-Lerma 2003c, Lem.6.2) and (Guo and Hernández-Lerma 2003a, Lem.3.1 (b)).

Lemma 2.6 *Suppose parts (a) and (b) of Conditions 2.1, Conditions 2.2, parts (b) and (c) of Conditions 2.3, as well as Conditions 2.4 (or Conditions 2.5) are satisfied. Then for any stationary policy $\pi(da|x)$, the following Kolmogorov's backward equation holds:*

$$\frac{dP_x^\pi(\xi_t \in \Gamma)}{dt} = \int_S P_y^\pi\{\xi_t \in \Gamma\} \int_A \pi(da|x) q(dy|x,a), \tag{2.17}$$

where $\Gamma \in \mathscr{B}(S)$; and the performance functional $V_0(x,\pi)$ satisfies

$$\alpha V_0(x,\pi) = \int_A \pi(da|x) c_0(x,a) + \int_S \int_A \pi(da|x) q(dy|x,a) V_0(y,\pi). \tag{2.18}$$

2.4 Convex analytic approach for constrained CTMDPs: $N > 0$

2.4.1 Occupation measure

Definition 2.3 *(a) The occupation measure of a policy π is a probability measure on $\mathscr{B}(S \times A)$ defined as follows:*

$$\eta^\pi(\Gamma) \triangleq \alpha \int_0^\infty e^{-\alpha t} \chi_t^\pi(\Gamma) dt, \tag{2.19}$$

for any $\Gamma \in \mathscr{B}(S \times A)$. Here for any $t \geq 0$, $\chi_t^\pi(dx,da)$ is the probability measure on $\mathscr{B}(S \times A)$ given by

$$\chi_t^\pi(\Gamma_S \times \Gamma_A) \triangleq E_\gamma^\pi [I\{\xi_t \in \Gamma_S\} \pi(\Gamma_A|\omega,t)],$$

where $\Gamma_S \in \mathscr{B}(S)$ and $\Gamma_A \in \mathscr{B}(A)$. Let $\mathscr{D}$ be the set of all occupation measures.
(b) Two policies π^1 and π^2 are said to be equivalent if $\eta^{\pi^1} = \eta^{\pi^2}$.

Note, we have chosen not to signify the dependence of occupation measures on γ and α for notational convenience.

Theorem 2.5 *Suppose parts (a) and (b) of Conditions 2.1, Conditions 2.2, parts (a) and (b) of Conditions 2.3, as well as Conditions 2.4 (or Conditions 2.5) are satisfied. Then the following assertions hold:*
(a) For any fixed π, η^π satisfies the following two relations:

$$\eta(\Gamma_S \times A) = \gamma(\Gamma_S) + \frac{1}{\alpha}\int_{S\times A} q(\Gamma_S|y,a)\eta(dy \times da), \tag{2.20}$$

where Γ_S is any set in $\mathscr{B}(S)$;

$$\int_S w(x)\eta(dx \times A) \le \frac{\alpha\int_S w(x)\gamma(dx) + b}{\alpha - \rho} < \infty, \tag{2.21}$$

where $w(x)$ comes from Conditions 2.1.
(b) Any occupation measure $\eta^\pi(dx \times da)$ is concentrated on K, in that $\eta^\pi(K) = 1$.
(c) If a (probability) measure on $S \times A$ concentrated on K, namely, η, satisfies the two relations in part (a), namely, (2.20) and (2.21), then there exists a stationary policy π such that $\eta = \eta^\pi$. Indeed, one can take π as in the following formula, whose validity is guaranteed by (Hernández-Lerma and Lasserre 1996, D.8 Prop.):

$$\eta(\Gamma_S \times \Gamma_A) = \int_{\Gamma_S} \pi(\Gamma_A|y)\eta(dy \times A), \tag{2.22}$$

where $\Gamma_S \in \mathscr{B}(S)$ and $\Gamma_A \in \mathscr{B}(A)$.
(d) Consider the probability measure η as in part (c). If a stationary policy $\hat{\pi}$ is such that $\eta^{\hat{\pi}} = \eta$, then $\hat{\pi}$ is a version of the π in (2.22), and in fact,

$$\eta^{\hat{\pi}}(\Gamma_S \times \Gamma_A) = \int_{\Gamma_S} \hat{\pi}(\Gamma_A|y)\eta^{\hat{\pi}}(dy \times A).$$

(e) Suppose a stationary policy π is fixed. Then equation

$$\eta(\Gamma_S) = \gamma(\Gamma_S) + \frac{1}{\alpha}\int_S\int_A q(\Gamma_S|y,a)\pi(da|y)\eta(dy), \tag{2.23}$$

where $\Gamma_S \in \mathscr{B}(S)$, has a unique solution in the class of probability measures on S subject to

$$\int_S w(x)\eta(dx) < \infty; \tag{2.24}$$

and the unique solution is provided by $\tilde{\eta}(dx) \triangleq \eta^\pi(dx \times A)$.

Remark 2.9 *(a) Theorem 2.5 implies that the space $\mathscr{D}$ is fully characterized by (probability) measures on $S \times A$, concentrated on K and satisfying relations (2.20) and (2.21). It then follows that under the conditions in Theorem 2.5, $\mathscr{D}$ is convex.*

(b) In what follows, under the conditions imposed in Theorem 2.5, conventionally, we will regard occupation measures η as measures on K. Consequently, $\forall\, \Gamma_S \in \mathscr{B}(S)$, by $\eta(\Gamma_S \times A)$ we mean $\eta((\Gamma_S \times A)\bigcap K)$, and so on. This is legal, because of part (b) of Theorem 2.5.

Definition 2.4 *Suppose $f(x) \geq 1$ is any given measurable function on S (and thus on K), and consider the linear space of finite signed measures M on K with a finite f-norm, where the latter means $\int_K f(x)|M|(dx \times da) < \infty$. This space is denoted as $\mathscr{M}_f^{\pm}(K)$, on which, let us define the f-weak convergence $M_n \xrightarrow{f} M$ taking place if for any continuous function $u(x,a)$ on K with a finite f-norm in that $\sup_{x\in S} \frac{\sup_{a\subset A(x)} |u(x,a)|}{f(x)} < \infty$, the following holds:*

$$\lim_{n\to\infty} \int_K u(x,a) M_n(dx \times da) = \int_K u(x,a) M(dx \times da).$$

Let us denote $\tau(\mathscr{M}_f^{\pm}(K))$ the weakest topology on $\mathscr{M}_f^{\pm}(K)$ such that $\int_K u(x,a)M(dx \times da)$, when viewed as a function on M, is continuous, for all u continuous on K and with a finite f-norm. Here we often refer this topology to as the f-weak topology.

The f-weak convergence of f-bounded signed measures on other Borel spaces can be defined similarly.

Let us call $\mathscr{M}_f(K)$ the set of all probability measures M on K such that $\int_K f(x)M(dx \times da) < \infty$, and $\mathscr{M}(K)$ the set of all probability measures $\tilde{M}$ on K. Then there exists a one-to-one correspondence between $M \in \mathscr{M}_f(K)$ and $\tilde{M} \in \mathscr{M}(K)$. Indeed, using any $M \in \mathscr{M}_f(K)$, one can define $\tilde{M} \in \mathscr{M}(K)$ via

$$\tilde{M}(\Gamma_S \times \Gamma_A) \triangleq \frac{\int_{\Gamma_S} f(x)M(dx \times \Gamma_A)}{\int_S f(x)M(dx \times A)}, \tag{2.25}$$

and using any $\tilde{M} \in \mathscr{M}(K)$, one can reproduce $M \in \mathscr{M}_f(K)$ via

$$M(\Gamma_S \times \Gamma_A) \triangleq \frac{\int_{\Gamma_S} \frac{\tilde{M}(dx \times \Gamma_A)}{f(x)}}{\int_S \frac{\tilde{M}(dx \times A)}{f(x)}}, \tag{2.26}$$

which can be vividly seen to provide the desired one-to-one correspondence. For notational convenience, we shall call Q_f this mapping from $\mathscr{M}_f(K)$ to $\mathscr{M}(K)$, and denote $\tilde{M}_n \to \tilde{M}$ for the usual weak convergence.

Remark 2.10 *If we call $\tilde{\rho}$ the metric on $\mathscr{M}(K)$ generating the usual weak topology (see, for example, Prokhorov's metric), then for any $M_1, M_2 \in \mathscr{M}_f(K)$, $\rho(M_1, M_2) \triangleq \tilde{\rho}(Q_f(M_1), Q_f(M_2))$ defines a metric on $\mathscr{M}_f(K)$. Clearly, Q_f and Q_f^{-1} are both continuous with respect to the metrics $\tilde{\rho}$ and ρ. Moreover, Lemma 2.7 below shows that if $f(x)$ is continuous on S, then the f-weak convergence stated in Definition 2.4 charac-*

terizes a topology on $\mathcal{M}_f(K)$, which is right generated by the above defined ρ. Indeed, following from (Gemignani 1990, Ex.7, p.127), this topology is the relative topology of $\tau(\mathcal{M}_f^{\pm}(K))$ restricted on $\mathcal{M}_f(K)$. Nevertheless, we shall often refer it to as f-weak topology, too, since the context will always exclude any confusion. Therefore, the two topological spaces $\mathcal{M}_f(K)$ and $\mathcal{M}(K)$ are homeomorphic.

Lemma 2.7 *Suppose some $f(x) \geq 1$ continuous on S has been fixed. Let M, $M_n \in \mathcal{M}_f(K)$, $n = 1, 2, \ldots$, and $\tilde{M}$, $\tilde{M}_n$ be the corresponding measures transformed via (2.25). Then $M_n \xrightarrow{f} M$ if and only if $\tilde{M}_n \to \tilde{M}$.*

Conditions 2.9 *(a) For any bounded continuous function u on S, $\int_S u(y)q(dy|x,a)$ is continuous on K.*
(b) w coming from Conditions 2.1 is continuous on S.

Part (a) of Conditions 2.9 says that $q(dy|x,a)$ is a weakly continuous transition rate (Hernández-Lerma and Lasserre 1999, Assump.8.5.1(c)), which is often assumed to prove the closeness of the space of occupation measures, see the condition in (Piunovskiy 1997, Thm.9, p.81). The condition requiring $w(x)$ to be continuous is practically not restrictive: the technique of using continuous functions to approximate an arbitrary function is well known, see for example, Luzin's theorem (Aliprantis and Border 2007, Thm.12.8) and its various generalizations.

Theorem 2.6 *Suppose parts (a) and (b) of Conditions 2.1, Conditions 2.2, parts (a) and (b) of Conditions 2.3, Conditions 2.4 (or Conditions 2.5), as well as Conditions 2.9 are satisfied. Then $\mathcal{D}$ is w-weakly closed (in $\mathcal{M}_w(K)$).*

2.4.2 Linear program formulation

Primal linear program

One approach of studying the optimization problem (2.13) is to formulate it as a linear program, which can then be hopefully solved directly or approximately. Clearly, according to Subsection 2.4.1, under parts (a) and (b) of Conditions 2.1, Conditions 2.2, Conditions 2.3, as well as Conditions 2.4 (or Conditions 2.5), the constrained optimization problem (2.13) is equivalent to the following PLP:

$$\frac{1}{\alpha}\int_K c_0(x,a)\eta(dx \times da) \to \min_{\eta,\beta_1,\ldots,\beta_N} \tag{2.27}$$

s.t.

$$\eta(\Gamma_S \times A) - \frac{1}{\alpha}\int_K q(\Gamma_S|x,a)\eta(dx \times da) = \gamma(\Gamma_S),\ \forall\, \Gamma_S \in \mathcal{B}(S);$$

$$\int_K c_n(x,a)\eta(dx\times da)+\beta_n=\alpha d_n,\ n=1,2,\ldots,N;$$
$$\int_S \eta(dx\times A)w(x)<\infty;$$
$$\eta \text{ is a probability measure on } K,\ \beta_n\geq 0,\ n=1,2,\ldots,N.$$

Indeed, for any $n=0,1,\ldots,N$, $V_n(\pi)=\frac{1}{\alpha}\int_{S\times A}\eta(dx\times da)c_n(x,a)<\infty$, by Conditions 2.3. Evidently, PLP (2.27) is feasible, due to Remark 2.1.

Sometimes, it is convenient to write PLP (2.27) in the following equivalent form:

$$\frac{1}{\alpha}\int_K c_0(x,a)\eta(dx\times da)\to\min_{\eta} \tag{2.28}$$

$$s.t.$$

$$\frac{1}{\alpha}\int_K c_n(x,a)\eta(dx\times da)-d_n\leq 0,\ n=1,2,\ldots,N;$$
$$\eta\in\mathscr{D},$$

where $\mathscr{D}$ is a nonempty convex set of probability measures on K, whose elements are fully characterized by (2.20) and (2.21). If additionally, Conditions 2.9 hold, then $\mathscr{D}$ is w-weakly closed, too, see Theorem 2.6.

In what follows, the optimal value of a PLP will be denoted as $\inf(PLP)$, and that of DLP will be treated similarly. We recall that here and below DLP means "Dual Linear Program".

Dual linear program

Sometimes, it appears more convenient to consider its associated DLP, after a PLP is formulated. However, in general, the optimal value of the DLP differs from that of the PLP, and the difference is often referred to as the duality gap. In below, after formulating the DLP of PLP (2.27), under some extra conditions, we shall show the absence of the duality gap.

Under parts (a) and (b) of Conditions 2.1, Conditions 2.2, Conditions 2.3, Conditions 2.4 (or Conditions 2.5) as well as parts (a), (b) and (c) of Conditions 2.6, let us formulate the following DLP in relation to PLP (2.27). To this end, according to (Hernández-Lerma and Lasserre 1996, Chap.6) (or (Hernández-Lerma and Lasserre 1999, Chap.12)), we need the following objects: two dual pairs $(\mathscr{X},\mathscr{Y})$ and $(\mathscr{Z},\mathscr{V})$, a positive cone Co in $\mathscr{X}$, a weakly continuous linear mapping from $\mathscr{X}$ to $\mathscr{Z}$, and two fixed points, namely, $B\in\mathscr{Z}$ and $C\in\mathscr{Y}$, which are collectively introduced in the below. In what follows, w and w' come from Conditions 2.1 and Conditions 2.6, respectively.

Let us denote $\mathscr{X}_0=\{\eta\}$ the linear space of all finite signed measures on K such that $\int_S w(x)|\eta|(dx\times A)<\infty$ (that is, $\mathscr{X}_0$ is taken as $\mathscr{M}_w^{\pm}(K)$), $\mathscr{Y}_0=\{f\}$ the linear space of all measurable functions on K such that $\sup_{x\in S}\frac{\sup_{a\in A(x)}|f(x,a)|}{w(x)}<\infty$, $\mathscr{Z}_0=\{z_0\}$ the linear space of finite signed measures on S such that $\int_S w'(x)|z_0|(dx)<\infty$, and $\mathscr{V}_0=$

$\{v'\} \triangleq \mathbf{B}_{w'}(S)$, which is clearly a linear space. Now, one can consider the following four linear spaces $\mathscr{X} \triangleq \mathscr{X}_0 \times \mathbb{R}^N = \{X = (\eta, \beta_1, \ldots, \beta_N)\}$, $\mathscr{Y} \triangleq \mathscr{Y}_0 \times \mathbb{R}^N = \{Y = (f, e_1, \ldots, e_N)\}$, $\mathscr{Z} \triangleq \mathscr{Z}_0 \times \mathbb{R}^N = \{Z = (z_0, h_1, \ldots, h_N)\}$ and $\mathscr{V} = \mathscr{V}_0 \times \mathbb{R}^N = \{V = (v', g'_1, \ldots, g'_N)\}$. Equipped with the bilinear forms $\langle X, Y\rangle \triangleq \int_K f(x,a)\eta(dx \times da) + \sum_{n=1}^N e_n\beta_n$ and $\langle Z, V\rangle \triangleq \int_S v'(x)z_0(dx) + \sum_{n=1}^N h_n g'_n$ respectively, we finally have the promised two dual pairs, namely $(\mathscr{X}, \mathscr{Y})$ and $(\mathscr{Z}, \mathscr{V})$. Here, space $\mathscr{X}$ is equipped with the weak topology $\tau(\mathscr{X}, \mathscr{Y})$, generated by all elements in $\mathscr{Y}$ when viewed as linear functionals on $\mathscr{X}$ through $\langle X, Y\rangle$, and similar assertions apply to the other three linear spaces under consideration. By the way, clearly, $\mathscr{M}_w(K)$ is w-weakly closed in $\mathscr{M}_w^{\pm}(K)$. Therefore, if $\mathscr{D}$ is w-weakly closed in $\mathscr{M}_w(K)$, it is also w-weakly closed in $\mathscr{M}_w^{\pm}(K)$. Consequently, $\mathscr{D} \times \mathbb{R}^n$ is closed in $\mathscr{X}_0 \times \mathbb{R}^n$ equipped with the product topology, where we have equipped $\mathscr{X}_0$ (resp. $\mathbb{R}^n$) with the w-weak topology (resp. the usual Euclidean topology). From this, it is not hard to see finally that the set $\mathscr{D} \times \mathbb{R}^n$ is also closed in $\tau(\mathscr{X}, \mathscr{Y})$, simply because this product topology on $\mathscr{X}_0 \times \mathbb{R}^n$ is contained in $\tau(\mathscr{X}, \mathscr{Y})$.

Next, let us define a linear mapping from $\mathscr{X}$ to $\mathscr{Z}$, namely, $Z = U \circ X$ by $z_0(\Gamma_S) = \eta(\Gamma_S \times A) - \frac{1}{\alpha}\int_K q(\Gamma_S|y,a)\eta(dy \times da)$, $h_n = \int_K c_n(x,a)\eta(dx \times da) + \beta_n$, where Γ_S is any set in $\mathscr{B}(S)$ and $n = 1, \ldots, N$. Then its adjoint mapping $Y = U^* \circ V$ is given by $f(x,a) = v'(x) - \frac{1}{\alpha}\int_S v'(y)q(dy|x,a) + \sum_{n=1}^N g'_n c_n(x,a)$, and $e_n = g'_n$, $n = 1, \ldots, N$, where clearly $f \in \mathscr{Y}_0$. Indeed,

$$
\begin{aligned}
\langle U \circ X, V\rangle &= \langle (z_0, h_1, \ldots, h_N), (v', g'_1, \ldots, g'_N)\rangle \\
&= \int_S v'(x)z_0(dx) + \sum_{n=1}^N h_n g'_n \\
&= \int_S v'(x)\left\{\eta(dx \times A) - \frac{1}{\alpha}\int_K q(dx|y,a)\eta(dy \times da)\right\} \\
&\quad + \sum_{n=1}^N \left(\int_K c_n(x,a)\eta(dx \times da) + \beta_n\right)g'_n \\
&= \int_K \eta(dx \times da)\left\{v'(x) - \frac{1}{\alpha}\int_S v'(y)q(dy|x,a) + \sum_{n=1}^N g'_n c_n(x,a)\right\} \\
&\quad + \sum_{n=1}^N \beta_n g'_n \\
&= \langle X, U^* \circ V\rangle,
\end{aligned}
$$

where the interchange of integrals are legal due to the imposed conditions. Now, one can infer from (Hernández-Lerma and Lasserre 1999, Prop.12.2.5) that U is the promised weakly continuous linear mapping from $\mathscr{X}$ to $\mathscr{Z}$.

Let us fix the following positive cone in $\mathscr{X}$, namely, $Co = \{\eta \geq 0, \beta_n \geq 0, n = 1, \ldots, N\}$. Evidently, its dual cone is given by $Co^* = \{f \geq 0, e_n \geq 0, n = 1, \ldots, N\}$.

Finally, we shall fix $B \triangleq (\gamma, \alpha d_1, \ldots, \alpha d_N) \in \mathscr{X}$ and $C \triangleq (\frac{1}{\alpha} c_0, 0, \ldots, 0) \in \mathscr{Y}$.

As promised, we shall now formulate our linear programs. According to (Hernández-Lerma and Lasserre 1996, Chap.6) (or (Hernández-Lerma and Lasserre 1999, Chap.12)), PLP (2.27) can be written concisely as

$$\begin{aligned} & \langle X, C \rangle \to \min_{X \in \mathscr{X}} \\ s.t. & \\ & U \circ X = B; \\ & X \in Co, \end{aligned} \tag{2.29}$$

and below is its dual program

$$\begin{aligned} & \langle B, V \rangle \to \max_{V \in \mathscr{V}} \\ s.t. & \\ & C - U^* \circ V \in Co^*; \\ & V \in \mathscr{V}, \end{aligned}$$

which can be written out more explicitly as follows:

$$\begin{aligned} & \int_S v'(x)\gamma(dx) + \alpha \sum_{n=1}^{N} d_n g'_n \to \max_{(v', g'_1, \ldots, g'_N) \in \mathscr{V}} \\ s.t. & \\ & \frac{1}{\alpha} c_0(x,a) - v'(x) + \frac{1}{\alpha} \int_S v'(y) q(dy|x,a) - \sum_{n=1}^{N} g'_n c_n(x,a) \geq 0; \\ & -g'_n \geq 0, \; n = 1, 2, \ldots N. \end{aligned}$$

However, it appears more convenient to rewrite the above DLP after the change of variables through $g_n \triangleq -g'_n \alpha$ and $v(x) \triangleq v'(x) - \sum_{n=1}^{N} d_n g_n$, and consequently we have

$$\begin{aligned} & \int_S \gamma(dx) v(x) \to \max_{(v, g_1, \ldots, g_N)} \\ s.t. & \\ & \frac{1}{\alpha} c_0(x,a) + \sum_{n=1}^{N} g_n \left(\frac{c_n(x,a)}{\alpha} - d_n\right) - v(x) + \frac{1}{\alpha} \int_S v(y) q(dy|x,a) \geq 0; \\ & g_n \geq 0, \; n = 1, \ldots, N; \\ & v \in \mathbf{B}_{w'}(S). \end{aligned} \tag{2.30}$$

If in addition, parts (b), (c) and (d) of Conditions 2.6 and Conditions 2.7 (or Conditions 2.8) are satisfied, then according to Remark 2.7, there is a function $u^* \in \mathbf{B}_{w'}(S)$ such that $\alpha u^*(x) \leq c_0(x,a) + \int_S u(y) q(dy|x,a)$, implying that $(u^*, 0, \ldots, 0)$ is a feasible solution to DLP (2.30). In other words, DLP (2.30) is consistent, which, together with the consistency of PLP (2.27) (see Subsection 2.4.2), leads to the weak duality, according to (Hernández-Lerma and Lasserre 1996, Thm.6.24). This is formulated in the following proposition:

Proposition 2.2 *Suppose parts (a) and (b) of Conditions 2.1, Conditions 2.2, Conditions 2.3, Conditions 2.4 (or Conditions 2.5), Conditions 2.6 and Conditions 2.7 (or Conditions 2.8) are satisfied. Then the following weak duality holds:*

$$-\infty < \sup(DLP\ (2.30)) \leq \inf(PLP\ (2.27)) < \infty.$$

Besides, if X is feasible for PLP (2.27), V is feasible for (2.30), as well as

$$\langle X, C - U^* \circ V \rangle = 0,$$

then X is optimal for PLP (2.27), and V is optimal for (2.30).

Conditions 2.10 *(a) For any bounded continuous function u on S, $\int_S u(y)q(dy|x,a)$ is continuous on K.*
(b) Functions w and w' are both continuous on S. Here we recall that w (resp. w') comes from Conditions 2.1 (resp. Conditions 2.6).

Theorem 2.7 *Suppose parts (a) and (b) of Conditions 2.1, Conditions 2.2, Conditions 2.3, Conditions 2.4 (or Conditions 2.5), Conditions 2.6, Conditions 2.7 (or Conditions 2.8), and Conditions 2.10 are satisfied. If additionally, the inequalities in (2.13) are all strict for some π (the so called Slater's condition), then there is no duality gap:*

$$-\infty < \sup(DLP\ (2.30)) = \inf(PLP\ (2.27)) < \infty.$$

2.4.3 Solvability of the constrained CTMDP

In below, we shall show the solvability of the constrained CTMDP optimization problem (2.13).

Conditions 2.11 *(a) For any bounded continuous function u on S, $\int_S u(y)q(dy|x,a)$ is continuous on K, and $\sup_{x\in S} \frac{\bar{q}_x}{w'(x)} < \infty$.*
(b) (i) The multifunction $x \rightarrow A(x)$ is compact-valued and upper semicontinuous (see (Hernández-Lerma and Lasserre 1996, Appendix D)).
(ii) w' is continuous, and $\forall\, r \geq 0$, $\exists$ a compact set $S_r \subseteq S$ such that $\frac{w(x)}{w'(x)} \geq r\ \forall\, x \notin S_r$.
(iii) S and A are σ-compact.
(c) $c_n(x,a)$ are lower semicontinuous on K.

Conditions 2.11 are a version of the so called compactness-continuity conditions, commonly assumed to prove the solvability of optimization problems, see for example, (Guo 2007, Assump.c), (Schäl 1975, Cons. (W) and (S)) and (Hernández-Lerma and Lasserre 1999, Assump.8.5.1, 8.5.2, 8.5.3). In particular, part (b) of Conditions 2.11 comes from (Hernández-Lerma and Lasserre 1996, Con.5.7.4, Rem.5.7.5).

Lemma 2.8 *Suppose parts (a) and (b) of Conditions 2.1, Conditions 2.2, parts (a) and (b) of Conditions 2.3, Conditions 2.4 (or Conditions 2.5), parts (a), (b) and (c) of Conditions 2.6 and parts (a) and (b) of Conditions 2.11 are satisfied. Then* $\mathscr{D}$ *is compact in the* w'*-weak topology.*

Theorem 2.8 *Suppose parts (a) and (b) of Conditions 2.1, Conditions 2.2, Conditions 2.3, Conditions 2.4 (or Conditions 2.5), Conditions 2.6 and Conditions 2.11 are satisfied. Then there is a stationary constrained-optimal policy to problem (2.13). Here we recall that* U*, the set of feasible policies, is not empty (see Remark 2.1).*

Note, in contrast with the case of unconstrained problems, it is well known that the class of deterministic stationary policies are not sufficient for solving constrained CTMDP optimization problems, see (Piunovskiy 1998, Sec.7).

2.5 Example

The following example is a modified version of the one in Guo (2007). Below, M_e, λ, μ_e p_0 and p_1 are fixed positive constants such that $\lambda M_e < 1$, and $Leb(dx)$ stands for the Lebesgue measure. Then the CTMDP admits the following the primitives:

- $S \triangleq [0,\infty)$,
- $A \triangleq [0,\infty)$,
- $A(x) \triangleq [0,x]$,
- $q(\Gamma_S|x,a) \triangleq \mu_e Leb(\Gamma_S \bigcap [x-a,x]) + \lambda Leb(\Gamma_S \bigcap [x,x+M_e]) - (\mu_e a + \lambda M_e) I\{x \in \Gamma_S\}$, $\forall\, \Gamma_S \in \mathscr{B}(S)$,
- $c_0(x,a) = p_0 x$,
- $c_1(x,a) = p_1 a$,
- $\alpha > \lambda M_e^2$,
- $\gamma(\cdot) \triangleq \delta_z(\cdot)$ with $z \in S$.

Now let us verify the validity of all the conditions required for our main theorems:
Conditions 2.1: Let us fix $w(x) \triangleq (1+x)^2$. Then parts (a) and (c) trivially hold, and as shown by Guo (2007), $\int_S w(y)q(dy|x,a) \leq \lambda M_e^2 w(x) + \lambda M_e^3$ validates part (b) with $\rho \triangleq \lambda M_e^2$ and $b \triangleq \lambda M_e^3$.

Conditions 2.2: As shown in Guo (2007), $\bar{q}_x \leq (\mu_e + \lambda M_e)w(x)$. That is, we only need put $L \triangleq \mu_e + \lambda M_e$.

Conditions 2.3: Parts (a) and (b) are trivially satisfied, whereas by fixing $M \triangleq p_0 + p_1$ and $c \triangleq 0$, part (c) follows from $|c_0(x,a)| = p_0 x$ and $|c_1(x,a)| = p_1 a \leq p_1 x$.

Conditions 2.4: Both parts hold trivially.

Conditions 2.6: Fixing $w'(x) \triangleq 1+x$, $L' \triangleq \mu_e + \lambda M_e + 1$, $\rho' \triangleq \frac{1}{2}\alpha$ and $b' \triangleq \lambda M_e^2$, part (a) follows from (Guo 2007, (19)), part (b) follows from (Guo 2007, (18)), and the remaining parts hold trivially.

Conditions 2.7: Part (a) holds because (Guo 2007, Assump.3(2)) does, as was shown therein. Part (b) follows from the trivial calculation $\int_S (1+y)^2 q(dy|x,a) = -\mu_e a^2 (x+1) + \frac{1}{3}\mu_e a^3 + \lambda M_e^2 (x+1) + \frac{1}{3}\lambda M_e^3$. The remaining parts hold trivially.

Conditions 2.9: For any bounded and continuous u on S, we have $\int_S u(y) q(dy|x,a) = \mu_e \int_{x-a}^{x} u(y)dy + \lambda \int_x^{x+M_e} u(y)dy - (\mu_e a + \Lambda M_e)u(x)$, following which and the continuity of u, we conclude that part (a) is satisfied, whereas part (b) holds obviously.

Conditions 2.10: They hold obviously.

Conditions 2.11: Parts (a) and (c) hold trivially, together with part (b) (ii, iii). To see why part (b) (i) holds, one only needs recall the definition of a upper semicontinuous multifunction, saying that for any F closed in $[0,\infty)$, $\{x \in S : [0,x] \bigcap F \neq \emptyset\}$ is closed (Hernández-Lerma and Lasserre 1996, Appendix D). Indeed, suppose $x_n \longrightarrow x$ as $n \longrightarrow \infty$ such that for each $n = 1,2,\ldots$, $[0,x_n] \bigcap F \neq \emptyset$. Due to Bolzano-Weierstrass' theorem, it suffices to consider when x_n converges to x monotonically, for otherwise, one simply needs take the corresponding subsequence. If $x_n \uparrow x$, then clearly $[0,x] \bigcap F \neq \emptyset$. Now consider the case $x_n \downarrow x$. Suppose $[0,x] \bigcap F = \emptyset$. Then for each n, there is a $y_n \in F$ such that $y \in (x, x_n]$. It follows from $x_n \downarrow x$ that $y_n \downarrow x$. However, $y_n \in F$ and F is closed, meaning that $x \in F$, which is a desired contradiction, as required.

Here we remind that Conditions 2.5 and Conditions 2.8 are not necessary for the validity of the above derived results, as soon as Conditions 2.4 and Conditions 2.7 hold. Suppose d_1 is sufficiently large. Then according to Corollary 2.1, obviously Slater's condition holds, and the optimization problem is feasible. Therefore, all the derived theorems in this chapter are applicable to this model, while, to our best knowledge, all the previous works on CTMDPs do not cover this constrained problem. Indeed, firstly, this optimization problem is a constrained one, meaning that the results (here by results we mean at least those concerning the solvability of the optimization problem) in Guo (2007); Yan et al (2008) are not applicable, because they are about unconstrained CTMDPs; secondly, the state space is Borel, meaning that results in Piunovskiy (2005); Guo and Hernández-Lerma (2003a); Guo and Piunovskiy (2010) are not applicable, since they are about the case of countable state space; and thirdly, the transition rate $\bar{q}_x = \mu_e x + \lambda M_e$ is unbounded, implying that the results in Piunovskiy (1998) are not applicable, as over there the transition rates are required to be uniformly bounded.

2.6 Conclusion

In conclusion, this chapter provided a first treatment to constrained discounted CTMDPs with the state space and action space being both Borel, and the transition and cost rates being possibly unbounded. After studying the dynamic programming approach, which led to the sufficiency of the class of deterministic stationary policies for solving unconstrained CTMDPs, we then applied the convex analytic approach to induce finally the existence of stationary constrained-optimal policies for the constrained CTMDPs, where the optimality is out of the class of history-dependent policies. Our conditions are mild, standard and commonly assumed in the current literature on CTMDPs with unbounded transition rates. The similar approach could be applied to the studies of CTMDPs with other criteria, too.

2.7 Proof of main statements

Proof of Lemma 2.1. Observe firstly that

$$\begin{aligned}
&\int_s^t \left\{ \exp(-\int_s^u F_x(v)dv) \int_{S\setminus\{x\}} f(dy|x,u)h(u,y,t) \right\} du + e^{-\int_s^t F_x(v)dv} w(x) \\
= &\int_s^t \exp(-\int_s^u F_x(v)dv) e^{\rho(t-u)} \left(\int_S f(dy|x,u)w(y) - f(\{x\}|x,u)w(x) \right) du \\
&+\frac{b}{\rho}\int_s^t \exp(-\int_s^u F_x(v)dv)e^{\rho(t-u)}F_x(u)du \\
&-\frac{b}{\rho}\int_s^t \exp(-\int_s^u F_x(v)dv)F_x(u)du + e^{-\int_s^t F_x(v)dv} w(x) \\
\leq &\int_s^t \exp(-\int_s^u F_x(v)dv)e^{\rho(t-u)} (\rho w(x) + b + F_x(u)w(x))\, du \\
&+\frac{b}{\rho}\int_s^t \exp(-\int_s^u F_x(v)dv)e^{\rho(t-u)}F_x(u)du \\
&-\frac{b}{\rho}\int_s^t \exp(-\int_s^u F_x(v)dv)F_x(u)du + e^{-\int_s^t F_x(v)dv} w(x).
\end{aligned}$$

The rest of the proof now becomes identical to that in the proof of (Guo and Hernández-Lerma 2003d, Lem.3.2(a), p.239). ■

Proof of Lemma 2.2. Let us start with some preliminaries facilitating this proof. Observe firstly that for any $l = 0,1,2,\ldots$,

$$g(dy|x,u) \triangleq \begin{cases} \Lambda^l(dy|x_0,\theta_1,x_1,\ldots,\theta_l,x,u) & \text{if } dy \bigcap \{x\} = \emptyset; \\ \\ -\Lambda^l(S|x_0,\theta_1,x_1,\ldots,\theta_l,x,u) & \text{if } dy = \{x\}, \end{cases}$$

where Λ^m has been defined in (2.3), is a signed kernel on $\mathscr{B}(S)$ given (x,u) and satisfies all the conditions in Lemma 2.1. Observe that

$$\begin{aligned}
\int_S g(dy|x,u)w(y) &= \int_{S\setminus\{x\}} g(dy|x,u)w(y) + g(\{x\}|x,u)w(x) \\
&= \int_{S\setminus\{x\}}\int_A \pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l)q(dy\setminus\{x\}|x,a)w(y) \\
&\quad -\int_A \pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l)q(S\setminus\{x\}|x,a)w(x) \\
&= \int_{S\setminus\{x\}}\int_A \pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l)q(dy|x,a)w(y) \\
&\quad +\int_A \pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l)q(\{x\}|x,a)w(x) \\
&= \int_S\int_A \pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l)q(dy|x,a)w(y) \\
&= \int_A\int_S q(dy|x,a)w(y)\pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l) \\
&\le \rho w(x)+b,
\end{aligned}$$

where the third equality is due to $q(dy|x,a)$ being conservative, and the last inequality is by Conditions 2.1 (b). Indeed, the change of orders of integrals as in the last equality will be used frequently in the below, and can be reasoned as follows:

$$\int_{S\setminus\{x\}}\int_A \pi(da|\omega,u+T_l)q(dy|x,a)w(y) = \int_A \pi(da|\omega,u+T_l)(\int_{S\setminus\{x\}} w(y)q(dy|a,x)).$$

Now one only needs combine $\int_A \pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l)(\int_{S\setminus\{x\}} w(y)q(dy|a,x))$ and $\int_A \pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l)q(\{x\}|x,a)w(x)$, with the former one being non-negative and the latter one being finite, to obtain

$$\begin{aligned}
&\int_S\int_A \pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l)q(dy|x,a)w(y) \\
&= \int_A \pi(da|x_0,\theta_1,x_1,\ldots,\theta_l,x,u+T_l)\int_S w(y)q(dy|a,x).
\end{aligned}$$

Therefore, by applying Lemma 2.1 to the above defined signed kernel g, we conclude that h given in (2.6) satisfies

$$\begin{aligned}
h(s,x,\tilde{t}) &= h(0,x,\tilde{t}-s) \\
&\ge \int_s^{\tilde{t}}\Big\{\exp(-\int_s^u \Lambda^l(S|x_0,\theta_1,x_1,\ldots,\theta_l,x,v)dv)
\end{aligned}$$

$$\int_S \Lambda^l(dy|x_0,\theta_1,x_1,\ldots,\theta_l,x,u)h(u,y,\tilde{t})\Big\}du$$
$$+\exp(-\int_s^{\tilde{t}} \Lambda^l(S|x_0,\theta_1,x_1,\ldots,\theta_l,x,v)dv)w(x) \qquad (2.31)$$

for each $x \in S$ and $0 \leq s \leq \tilde{t} < \infty$.

As for the statement in Lemma 2.2, firstly, let us consider the case of $\rho > 0$. In fact, we shall prove inductively a slightly stronger result: fixing arbitrarily some m and x, for $n = 0,1,\ldots,m$

$$\begin{aligned} E_x^\pi\left[w(\xi_t)I\{t < T_{m+1}\}|\mathscr{F}_{T_{m-n}}\right] &\leq I\{T_{m-n} \leq t\}h(T_{m-n},x_{m-n},t) \\ &\quad + I\{T_{m-n} > t\}\sum_{k=1}^{m-n} I\{T_{k-1} \leq t < T_k\}w(x_{k-1}), \end{aligned}$$

where $\mathscr{F}_{T_n} \triangleq \sigma(x_m,T_m : 0 \leq m \leq n)$. We shall refer this statement to as a "stronger" statement to distinguish it from that in Lemma 2.2.

Step 1. Consider the case of $n = 0$:

$$\begin{aligned} E_x^\pi\left[w(\xi_t)I\{t < T_{m+1}\}|\mathscr{F}_{T_m}\right] &= E_x^\pi\left[(I\{T_m \leq t\} + I\{T_m > t\})w(\xi_t)\right. \\ &\quad \left.\times I\{t < T_{m+1}\}|\mathscr{F}_{T_m}\right] \\ &= I\{T_m \leq t\}w(x_m)P_x^\pi(\theta_{m+1} > t - T_m|\mathscr{F}_{T_m}) \\ &\quad + I\{T_m > t\}\sum_{k=1}^{m} I\{T_{k-1} \leq t < T_k\}w(x_{k-1}) \\ &= I\{T_m \leq t\}w(x_m)\exp(-\int_0^{t-T_m} \Lambda^m(S|h_m,v)dv) \\ &\quad + I\{T_m > t\}\sum_{k=1}^{m} I\{T_{k-1} \leq t < T_k\}w(x_{k-1}), \end{aligned}$$

where the last equality follows from the fact that on the set $\{T_m \leq t\}$, (2.4) implies

$$P_x^\pi(\theta_{m+1} > t - T_m|\mathscr{F}_{T_m}) = \exp(-\int_0^{t-T_m} \Lambda^m(S|h_m,v)dv). \qquad (2.32)$$

Due to (2.31), by estimating the first term in the last expression from the above, we have

$$\begin{aligned} E_x^\pi\left[w(\xi_t)I\{t < T_{m+1}\}|\mathscr{F}_{T_m}\right] &\leq I\{T_m \leq t\}h(T_m,x_m,t) \\ &\quad + I\{T_m > t\}\sum_{k=1}^{m} I\{T_{k-1} \leq t < T_k\}w(x_{k-1}). \end{aligned}$$

Step 2. Now suppose for $0 \leq n < m$, the "stronger" statement holds.

Step 3. Then consider the case of $n+1$:

$$\begin{aligned}
& E_x^\pi\left[w(\xi_t)I\{t<T_{m+1}\}|\mathscr{F}_{T_{m-n-1}}\right] \\
= \; & E_x^\pi\left[E_x^\pi\left[w(\xi_t)I\{t<T_{m+1}\}|\mathscr{F}_{T_{m-n}}\right]|\mathscr{F}_{T_{m-n-1}}\right] \\
\leq \; & E_x^\pi\left[I\{T_{m-n}\leq t\}h(T_{m-n},x_{m-n},t)+I\{T_{m-n}>t\}\sum_{k=1}^{m-n}I\{T_{k-1}\leq t<T_k\}\right. \\
& \left.\times w(x_{k-1})|\mathscr{F}_{T_{m-n-1}}\right] \\
= \; & E_x^\pi\left[I\{T_{m-n-1}\leq t\}I\{T_{m-n}\leq t\}h(T_{m-n},x_{m-n},t)|\mathscr{F}_{T_{m-n-1}}\right] \\
& +E_x^\pi\left[I\{T_{m-n-1}\leq t\}I\{T_{m-n}>t\}\sum_{k=1}^{m-n}I\{T_{k-1}\leq t<T_k\}w(x_{k-1})|\mathscr{F}_{T_{m-n-1}}\right] \\
& +E_x^\pi\left[I\{T_{m-n-1}>t\}\sum_{k=1}^{m-n}I\{T_{k-1}\leq t<T_k\}w(x_{k-1})|\mathscr{F}_{T_{m-n-1}}\right] \\
= \; & I\{T_{m-n-1}\leq t\}\left\{\int_0^{t-T_{m-n-1}}\left\{\exp(-\int_0^u\Lambda^{m-n-1}(S|h_{m-n-1},v)dv)\right.\right. \\
& \left.\times\int_{S\setminus\{x_{m-n-1}\}}\Lambda^{m-n-1}(dy|h_{m-n-1},u)h(T_{m-n-1}+u,y,t)\right\}du \\
& \left.+\exp(-\int_0^{t-T_{m-n-1}}\Lambda^{m-n-1}(S|h_{m-n-1},v)dv)w(x_{m-n-1})\right\} \\
& +I\{T_{m-n-1}>t\}\sum_{k=1}^{m-n-1}I\{T_{k-1}\leq t<T_k\}w(x_{k-1}),
\end{aligned}$$

where the first inequality follows from the inductive supposition, and the last equality from (2.32). Moreover, applying (2.31) to the term inside the parenthesis in the above derived expression gives

$$\begin{aligned}
E_x^\pi\left[w(\xi_t)I\{t<T_{m+1}\}|\mathscr{F}_{T_{m-n-1}}\right] \;\leq\; & I\{T_{m-n-1}\leq t\}h(T_{m-n-1},x_{m-n-1},t) \\
& +I\{T_{m-n-1}>t\}\sum_{k=1}^{m-n-1}I\{T_{k-1}\leq t<T_k\} \\
& \times w(x_{k-1}).
\end{aligned}$$

Hence, the "stronger" statement holds. It only remains to put $n=m$ in the "stronger" statement to obtain Lemma 2.2 for the case of $\rho>0$.

The case of $\rho=0$ follows now: one only needs observe that $\lim_{\rho\downarrow 0}\{e^{\rho t}w(x)+\frac{b}{\rho}(e^{\rho t}-1)\}=w(x)+bt.$ ■

Proof of Theorem 2.1. (a) In what follows, we shall quite formally regard $S_{-1}\triangleq\{x_\infty\}$. Preliminarily, let us prove first of all the following statement

$$\lim_{l\to\infty}P_x^\pi(\Gamma_l)=0, \tag{2.33}$$

where $\Gamma_l \triangleq \{\exists\, m : (\xi_t \in S\setminus S_l)\bigcap(T_m \le t < T_{m+1})\}$.

Before proving (2.33), observe that due to Conditions 2.1, $\forall\ \varepsilon > 0$, $\exists\ J(\varepsilon) > 0 :$ $\forall\ l \ge J(\varepsilon)$, we have

$$\inf_{y\in S\setminus S_l} w(y) > \frac{e^{\tilde{\rho}t}w(x) + \frac{b}{\tilde{\rho}}(e^{\tilde{\rho}t} - 1)}{\varepsilon}, \tag{2.34}$$

where $\tilde{\rho} \triangleq \rho + 1$.

Suppose now on the opposite that (2.33) does not hold. That is, $\exists\ \varepsilon > 0 : \forall\ L > 0$ $\exists\ l \ge \max\{L,\ J(\varepsilon)\}$ such that

$$P_x^\pi(\Gamma_l) > \varepsilon. \tag{2.35}$$

Necessarily (2.34) holds as well.

Let us modify the transition rates as follows:

$$\tilde{q}(dy|x,a) \triangleq \begin{cases} q(dy|x,a), & \text{if } x \in S_l; \\ 0, & \text{if } x \in S\setminus S_l. \end{cases}$$

The corresponding probabilities and expectations are equipped with the "tilde". Clearly, it follows (see also (2.4)) that

$$P_x^\pi(\forall\ m, (\xi_t \in S_l)\bigcap(T_m \le t < T_{m+1})) = \tilde{P}_x^\pi(\forall\ m, (\xi_t \in S_l)\bigcap(T_m \le t < T_{m+1})),$$

which, together with (2.35), implies

$$\tilde{P}_x^\pi(\Gamma_l) = P_x^\pi(\Gamma_l) > \varepsilon. \tag{2.36}$$

Evidently, if Conditions 2.1 hold for ρ and q, then they also hold for $\tilde{\rho}$ and $\tilde{q}$. Therefore, one can still apply Lemma 2.2 to induce $\forall\ m$

$$\tilde{E}_x^\pi\left[w(\xi_t)I\{t < T_{m+1}\}\right] \le e^{\tilde{\rho}t}w(x) + \frac{b}{\tilde{\rho}}(e^{\tilde{\rho}t} - 1),$$

which implies

$$\begin{aligned} \tilde{E}_x^\pi\left[w(\xi_t)\right] &= \tilde{E}_x^\pi\left[w(\xi_t)\sum_{m=0}^{\infty} I\{T_m \le t < T_{m+1}\}\right] \\ &\le e^{\tilde{\rho}t}w(x) + \frac{b}{\tilde{\rho}}(e^{\tilde{\rho}t} - 1), \end{aligned} \tag{2.37}$$

where the equality follows from the fact that $\sum_{m=0}^{\infty}\tilde{P}_x^\pi(T_m \le t < T_{m+1}) = 1$, which in turn is a result of $\sup_{x\in S}\sup_{a\in A(x)}\tilde{q}_x(a) \le \sup_{x\in S_l}\bar{q}_x < \infty$ (see Conditions 2.1).

On the other hand, we have

$$\begin{aligned}\tilde{E}_x^\pi[w(\xi_t)] &= \tilde{E}_x^\pi[w(\xi_t)|\Gamma_l]\tilde{P}_x^\pi(\Gamma_l) + \tilde{E}_x^\pi[w(\xi_t)|\Gamma_l^c]\tilde{P}_x^\pi(\Gamma_l^c)\\ &> \inf_{y\in S\setminus S_l} w(y)\varepsilon > e^{\tilde{\rho}t}w(x) + \frac{b}{\tilde{\rho}}(e^{\tilde{\rho}t}-1),\end{aligned}$$

where the first inequality follows from ignoring the second term in the first line and estimating the first term from below using (2.36), and the last inequality is a result of (2.34). However, this contradicts (2.37)! Hence, the statement (2.33) has been proved.

Let us now prove statement (a) in the theorem. Due to $\Gamma_{l+1} \subseteq \Gamma_l$ and (2.33), we conclude that $P_x^\pi(\bigcap_{l\ge 0}\Gamma_l) = 0$. In other words,

$$\forall\,\pi, x \in S, t \ge 0,\ P_x^\pi(\forall\, l, \exists\, m, \xi_t \in S\setminus S_l, T_m \le t < T_{m+1}) = 0. \tag{2.38}$$

Since $\{\inf\{s : \xi_s \in S\setminus S_l\} \le t\} \subseteq \{\exists\, m : \xi_t \in S\setminus S_l, T_m \le t < T_{m+1}\}$, with in mind (2.38), we have $P_x^\pi(\forall\, l : \inf\{s : \xi_s \in S\setminus S_l\} \le t) = 0$ and $P_x^\pi(\exists\, l : \forall\, s \in [0,t], \xi_s \in S_l) = 1$. But if $\xi_s \in S_l$, then Conditions 2.1 imply that $\bar{q}_{\xi_s}$ is going to be uniformly bounded, under which $T_\infty > t$, or $P_x^\pi(T_\infty > t) = 1$. Since t is arbitrary, this leads to $P_x^\pi(T_\infty = \infty) = 1$, as required.

(b) According to part (a), we have $\forall\, t$,

$$\sum_{m=0}^{\infty} P_x^\pi(T_m \le t < T_{m+1}) = 1. \tag{2.39}$$

Thus,

$$E_x^\pi[w(\xi_t)] = E_x^\pi\left[w(\xi_t)\sum_{m=0}^{\infty} I\{T_m \le t < T_{m+1}\}\right] = \lim_{m\to\infty} E_x^\pi[w(\xi_t)I\{t < T_{m+1}\}].$$

It only remains now to apply Lemma 2.2 to the above derived expression to induce the statement.

(c) Evidently,

$$\begin{aligned}P_x^\pi(\xi_t \in S) &= \sum_{m\ge 0} P_x^\pi(\xi_t \in S|T_m \le t < T_{m+1})P_x^\pi(T_m \le t < T_{m+1})\\ &= \sum_{m\ge 0} P_x^\pi(T_m \le t < T_{m+1}).\end{aligned}$$

It only remains now to recognize and apply (2.39). ■

Proof of Corollary 2.1. Consider firstly $|V_n(x,\pi)|$. Clearly,

$$\begin{aligned}
|V_n(x,\pi)| &\leq E_x^\pi\left[\int_0^\infty e^{-\alpha t}\int_A |c_n(\xi_{t-},a)|\pi(da|\omega,t)dt\right]\\
&\leq E_x^\pi\left[\int_0^\infty e^{-\alpha t}M(w(\xi_{t-})+c)dt\right]\\
&= \int_0^\infty e^{-\alpha t}(ME_x^\pi[w(\xi_{t-})]+c)dt\\
&\leq \int_0^\infty e^{-\alpha t}(M(e^{\rho t}w(x)+\frac{b}{\rho}(e^{\rho t}-1))+c)dt\\
&= \frac{M(\alpha w(x)+b)}{\alpha(\alpha-\rho)}+\frac{c}{\alpha},
\end{aligned}$$

where the second inequality follows from Conditions 2.3 (c), the last inequality from Theorem 2.1, and the final equality follows from Conditions 2.3 (b).

Now $|V_n(\pi)| \leq \frac{M(\alpha\int_S \gamma(dy)w(y)+b)}{\alpha(\alpha-\rho)}+\frac{c}{\alpha}<\infty$ follows, with in mind Conditions 2.3 (a).
■

Proof of Theorem 2.2. Define the following two random measures similar to μ and ν:

$$\tilde{\mu}(\omega,dt,\Gamma) \triangleq \sum_{m\geq 1} I\{T_m<\infty\}I\{x_{m-1}\in\Gamma\}I\{T_m\in dt\}$$

and

$$\tilde{\nu}(\omega,dt,\Gamma) \triangleq \int_A \pi(da|\omega,t)q(S\setminus\{\xi_{t-}\}|\xi_{t-},a)I\{\xi_{t-}\in\Gamma\}dt,$$

where $\Gamma\in\mathscr{B}(S)$. Then it was shown that $\tilde{\nu}$ is the dual predictable projection of $\tilde{\mu}$ (see the proof of (Kitaev 1986, Lem.4)) in that for any non-negative $\mathscr{P}\times\mathscr{B}(S)$ (the product σ-algebra)- measurable function $\kappa(\omega,t,x)$, the following relation holds:

$$E_x^\pi\left[\int_0^\infty\int_S \tilde{\mu}(dt,dy)\kappa(t,y)\right]=E_x^\pi\left[\int_0^\infty\int_S \tilde{\nu}(dt,dy)\kappa(t,y)\right],$$

see (Kitaev and Rykov 1995, Chap.4.5), following from which, one has

$$\begin{aligned}
E_x^\pi[\tilde{\mu}((0,t],\Gamma)] &= E_x^\pi[\tilde{\nu}((0,t],\Gamma)]=E_x^\pi\left[\int_0^t\int_A \pi(da|\omega,u)q_{\xi_{u-}}(a)I\{\xi_{u-}\in\Gamma\}du\right]\\
&\leq E_x^\pi\left[\int_0^t \bar{q}_{\xi_{u-}}I\{\xi_{u-}\in\Gamma\}du\right]\leq t\sup_{y\in S_l}\bar{q}_y<\infty, \qquad (2.40)
\end{aligned}$$

where the second inequality follows from part (c) of Conditions 2.1.

On the other hand, due to Theorem 2.1, μ and $\tilde{\mu}$ are both finite (a.s.), and furthermore, it follows from their definitions that $|\mu((0,t],\Gamma)-\tilde{\mu}((0,t],\Gamma)|\leq 1$, (a.s.). Now, it is obvious that (2.40) implies

$$E_x^\pi[\mu((0,t],\Gamma)]<\infty$$

as well.

Finally, the statement follows from taking the expectation legally in both sides of the following obvious expression

$$I\{\xi_t \in \Gamma\} = I\{\xi_0 \in \Gamma\} + \mu((0,t],\Gamma) - \tilde{\mu}((0,t],\Gamma),$$

which itself holds because of part (a) of Theorem 2.1. ■

Proof of Theorem 2.3. Let us first of all prove (2.11) for functions $r(x)$ of the form $r(x) \triangleq u(x)I\{x \in S_l\}$, with S_l as in Conditions 2.1. Observe firstly that the statement would follow immediately, if one can integrate $r(x)$ over S with respect to $P_x^\pi(\xi_t \in dy)$, within mind the Kolmogorov's forward equation derived in Theorem 2.2. Therefore, one only needs verify

$$\int_S r(y) E_x^\pi \left[\int_0^t \int_A \pi(da|\omega,v) q(dy \setminus \{\xi_v\}) |\xi_v, a) dv \right] < \infty \tag{2.41}$$

and

$$E_x^\pi \left[\int_0^t \int_A \pi(da|\omega,v) q_{\xi_v}(a) r(\xi_v) dv \right] < \infty. \tag{2.42}$$

Let us verify (2.41) now.

$$\begin{aligned}
& \int_S r(y) E_x^\pi \left[\int_0^t \int_A \pi(da|\omega,v) q(dy \setminus \{\xi_v\} |\xi_v, a) dv \right] \\
\leq\ & ||r||_{w'} \int_S w'(y) E_x^\pi \left[\int_0^t \int_A \pi(da|\omega,v) q(dy \setminus \{\xi_v\} |\xi_v, a) dv \right],
\end{aligned}$$

where

$$\begin{aligned}
& \int_S w'(y) E_x^\pi \left[\int_0^t \int_A \pi(da|\omega,v) q(dy \setminus \{\xi_v\} |\xi_v, a) dv \right] \\
=\ & E_x^\pi \left[\int_0^t \int_A \pi(da|\omega,v) \int_S w'(y) q(dy \setminus \{\xi_v\} |\xi_v, a) dv \right] \\
=\ & E_x^\pi \left[\int_0^t \int_A \pi(da|\omega,v) \int_S w'(y) \{q(dy|\xi_v, a) - q(\{\xi_v\}|\xi_v, a) I\{\xi_v \in dy\}\} dv \right].
\end{aligned}$$

Observe firstly that

$$\begin{aligned}
& E_x^\pi \left[\int_0^t \int_A \pi(da|\omega,v) \int_S w'(y) q(dy|\xi_v, a) dv \right] \\
\leq\ & E_x^\pi \left[\int_0^t \int_A \pi(da|\omega,v)(\rho' w'(\xi_v) + b') dv \right] \\
\leq\ & L'\rho' \int_0^t E_x^\pi [w(\xi_v)] dv + b't < \infty,
\end{aligned}$$

where the first inequality follows from Conditions 2.6 (b), the second inequality follows from Conditions 2.6 (a), and the final inequality follows from Theorem 2.1; and secondly that

$$
\begin{aligned}
& -E_x^\pi\left[\int_0^t \int_A \pi(da|\omega,v) w'(\xi_v) q(\{\xi_v\}|\xi_v,a)\right] \\
= \; & E_x^\pi\left[\int_0^t \int_A \pi(da|\omega,v) w'(\xi_v) q_{\xi_v}(a) dv\right] \le \int_0^t E_x^\pi\left[\bar{q}_{\xi_v} w'(\xi_v)\right] dv \\
\le \; & L' \int_0^t E_x^\pi\left[w(\xi_v)\right] dv < \infty,
\end{aligned}
$$

where the second inequality follows from Conditions 2.6(a), and the final inequality follows from Theorem 2.1. Therefore, we have verified (2.41).

In the process of verifying (2.41), (2.42) has been incidentally verified. Thus (2.11) is proved for $r(x)$.

Let us now prove (2.11) for $u(x) \in \mathbf{B}_{w'}(S)$. In what follows, we shall put $S_{-1} \triangleq \emptyset$. Then we have

$$
\begin{aligned}
& E_x^\pi\left[u(\xi_t)\right] - u(x) \\
= \; & E_x^\pi\left[\sum_{l=-1}^{\infty} u(\xi_t) I\{\xi_t \in S_{l+1} \setminus S_l\}\right] - \sum_{l=-1}^{\infty} u(x) I\{x \in S_{l+1} \setminus S_l\} \\
= \; & \sum_{l=-1}^{\infty} E_x^\pi\left[u(\xi_t) I\{\xi_t \in S_{l+1} \setminus S_l\}\right] - \sum_{l=-1}^{\infty} u(x) I\{x \in S_{l+1} \setminus S_l\} \\
= \; & \sum_{l=-1}^{\infty} \left\{E_x^\pi\left[u(\xi_t) I\{\xi_t \in S_{l+1} \setminus S_l\}\right] - u(x) I\{x \in S_{l+1} \setminus S_l\}\right\},
\end{aligned}
$$

where the interchange of integrals and expectations in the third line follows from that $E_x^\pi\left[\sum_{l=-1}^{\infty} |u(\xi_t)| I\{\xi_t \in S_{l+1} \setminus S_l\}\right] < \infty$.

Since it has been shown that (2.11) holds for $r(x)$ in the beginning of this proof, following from the obtained expression in the above, we have

$$
\begin{aligned}
& E_x^\pi\left[u(\xi_t)\right] - u(x) \\
= \; & \sum_{l=-1}^{\infty} \left\{E_x^\pi\left[\int_0^t \int_S \int_A \pi(da|\omega,v) q(dy \setminus (\xi_v)|\xi_v,a) u(y) I\{y \in S_{l+1} \setminus S_l\}\right]\right. \\
& \left. -E_x^\pi\left[\int_0^t \int_A \pi(da|\omega,v) q_{\xi_v}(a) u(\xi_v) I\{\xi_v \in S_{l+1} \setminus S_l\} dv\right]\right\}.
\end{aligned}
$$

Since

$$
\begin{aligned}
& \sum_{l=-1}^{\infty} \left\{E_x^\pi\left[\int_0^t \int_S \int_A \pi(da|\omega,v) q(dy \setminus (\xi_v)|\xi_v,a) u(y) I\{y \in S_{l+1} \setminus S_l\}\right]\right\} \\
= \; & E_x^\pi\left[\int_0^t \int_S \int_A \pi(da|\omega,v) q(dy \setminus (\xi_v)|\xi_v,a) u(y)\right] < \infty,
\end{aligned}
$$

which can be verified similarly as for (2.41) in the above for $r(x)$, and analogously

$$\begin{aligned}
&\sum_{l=-1}^{\infty}\left\{E_x^\pi\left[\int_0^t\int_A \pi(da|\omega,v)q_{\xi_v}(a)u(\xi_v)I\{\xi_v\in S_{l+1}\setminus S_l\}dv\right]\right\}\\
=\;& E_x^\pi\left[\int_0^t\int_A \pi(da|\omega,v)q_{\xi_v}(a)u(\xi_v)dv\right]<\infty,
\end{aligned}$$

we eventually conclude that

$$E_x^\pi[u(\xi_t)]-u(x)=E_x^\pi\left[\int_0^t\int_S\int_A \pi(da|\omega,v)q(dy|\xi_v,a)u(y)dv\right],$$

as required.

Let us prove (2.12) now. In what follows we shall repeatedly apply (2.11) to $E_x^\pi[u(\xi_t)]$.

On one hand, we have

$$\begin{aligned}
\text{LHS of (2.12)} &= e^{-\alpha t}\left\{u(x)+E_x^\pi\left[\int_0^t\int_S\int_A \pi(da|\omega,v)q(dy|\xi_v,a)u(y)dv\right]\right\}-u(x)\\
&= e^{-\alpha t}E_x^\pi\left[\int_0^t\int_S\int_A \pi(da|\omega,v)q(dy|\xi_v,a)u(y)dv\right]+u(x)(e^{-\alpha t}-1).
\end{aligned}$$

On the other hand, we observe firstly

$$\begin{aligned}
& E_x^\pi\left[\int_0^t e^{-\alpha v}(-\alpha u(\xi_v))dv\right]\\
=\;& -\alpha\int_0^t e^{-\alpha v}E_x^\pi[u(\xi_v)]\,dv\\
=\;& -\alpha\int_0^t e^{-\alpha v}\left\{u(x)+E_x^\pi\left[\int_0^v\int_S\int_A \pi(da|\omega,r)q(dy|\xi_r,a)u(y)dr\right]\right\}dv\\
=\;& (e^{-\alpha t}-1)u(x)-\alpha\int_0^t e^{-\alpha v}E_x^\pi\left[\int_0^v\int_S\int_A \pi(da|\omega,r)q(dy|\xi_r,a)u(y)dr\right]dv\\
=\;& (e^{-\alpha t}-1)u(x)-\alpha E_x^\pi\left[\int_0^t\left\{e^{-\alpha v}\int_0^v\int_S\int_A \pi(da|\omega,r)q(dy|\xi_r,a)u(y)dr\right\}dv\right]
\end{aligned}$$

where the interchange of integrals in the first and the last equalities is legal, because evidently for $u\in\mathbf{B}_{w'}(S)$,

$$E_x^\pi\left[\int_0^t e^{-\alpha v}\alpha u(\xi_v)dv\right]<\infty$$

as well as

$$\int_0^t e^{-\alpha v}E_x^\pi\left[\int_0^v\int_S\int_A \pi(da|\omega,r)q(dy|\xi_r,a)u(y)dr\right]dv<\infty$$

(see the proof of (2.11) for example); and secondly

$$\begin{aligned} & E_x^\pi\left[\int_0^t e^{-\alpha v}\int_S\int_A \pi(da|\omega,v)q(dy|\xi_v,a)u(y)dv\right] \\ = & E_x^\pi\left[e^{-\alpha t}\int_0^t\int_S\int_A \pi(da|\omega,r)q(dy|\xi_r,a)u(y)dr\right] \\ & +\alpha E_x^\pi\left[\int_0^t e^{-\alpha v}\int_0^v\int_S\int_A \pi(da|\omega,r)q(dy|\xi_r,a)u(y)dr\right], \end{aligned}$$

which altogether amount to

$$\begin{aligned} \text{RHS of (2.12)} &= E_x^\pi\left[\int_0^t e^{-\alpha v}(-\alpha u(\xi_v))dv\right] \\ & +E_x^\pi\left[\int_0^t e^{-\alpha v}\int_S\int_A \pi(da|\omega,v)q(dy|\xi_v,a)u(y)dv\right] \\ &= (e^{-\alpha t}-1)u(x) \\ & +E_x^\pi\left[e^{-\alpha t}\int_0^t\int_S\int_A \pi(da|\omega,r)q(dy|\xi_r,a)u(y)dr\right] \\ &= \text{LHS of (2.12)}, \end{aligned}$$

as required. ■

Proof of Lemma 2.4. Within mind of the remark immediately after Conditions 2.7, by the virtue of (Hernández-Lerma and Lasserre 1999, Lem.8.3.7(a)), under Conditions 2.7 (a) and (b), for any $u\in\mathbf{B}_w(S)$, $\int_S u(y)\left(\frac{q(dy|x,a)}{1+\bar{q}_x}+I\{x\in dy\}\right)$ is continuous in $a\in A(x)$ for any fixed $x\in S$. This, in addition to Conditions 2.7 (c), implies that the term inside the parenthesis is lower semicontinuous in $a\in A(x)$ for any fixed $x\in S$. On the other hand, under Conditions 2.7, due to the remark immediately after Conditions 2.7, we have that $\bar{q}_x$ is measurable on S. Therefore, by the virtue of (Bertsekas and Shreve 1978, Prop.7.29), the term in the parenthesis is also measurable on K. With in mind of Conditions 2.4 (a), it only remains now to apply (Hernández-Lerma and Lasserre 1996, D.5 Prop.) (see also (Bertsekas and Shreve 1978, Prop.7.33)). ■

Proof of Lemma 2.5. Let us firstly understand the roles played by the imposed conditions in the statement. Conditions 2.7 (or Conditions 2.8) guarantee the measurability of the regarding functions, based on which Conditions 2.3 (a) and (b) together with Conditions 2.6 and Conditions 2.1, which validate Theorem 2.1, ensure that all the respective integrals and expectations in the below are finite, and thus legalize operations such as passing limit inside integrals; according to Remark 2.7, Conditions 2.6 (b), (c) and (d) imply $u^*\in\mathbf{B}_{w'}(S)$; according to Theorem 2.3, Conditions 2.6 (a) and (b) as well as Conditions 2.1 validate the Dynkin's formula (2.12).

Under the conditions imposed in the statement, the Dynkin's formula (2.12) leads to that

$$e^{-\alpha t}E_\gamma^\pi[u^*(\xi_t)] = \int_S \gamma(dy)u^*(y) + E_\gamma^\pi\left[\int_0^t e^{-\alpha v}\int_A \pi(da|\omega,v)\{-\alpha u^*(\xi_v) + \int_S q(dy|\xi_v,a)u^*(y)\right\}dv\right]. \tag{2.43}$$

Adding the evidently finite expression $E_\gamma^\pi\left[\int_0^t e^{-\alpha v}\int_A \pi(da|\omega,v)c_0(\xi_v,a)dv\right]$ to the both sides of (2.43), we have

$$E_\gamma^\pi\left[\int_0^t e^{-\alpha v}\int_A \pi(da|\omega,v)c_0(\xi_v,a)dv\right] + e^{-\alpha t}E_\gamma^\pi[u^*(\xi_t)]$$

$$= \int_S \gamma(dy)u^*(y) + E_\gamma^\pi\left[\int_0^t e^{-\alpha v}\int_A \pi(da|\omega,v)\{c_0(\xi_v,a) - \alpha u^*(\xi_v) + \int_S q(dy|\xi_v,a)u^*(y)\right\}dv\right].$$

It only remains now to take the limit as $t \to \infty$ quite formally in both sides of the above derived expressions. Remember, here $\lim_{t\to\infty} e^{-\alpha t}E_\gamma^\pi[u^*(\xi_t)] = 0$, as a result of Theorem 2.1. ■

Proof of Theorem 2.4. (a) Since u^* solves the Bellman equation (2.14), according to (Hernández-Lerma and Lasserre 1996, D.5 Prop.) in case Conditions 2.7 are satisfied (or according to (Bertsekas and Shreve 1978, Prop.7.34) in case Conditions 2.8 are satisfied), for any $\varepsilon > 0$, one can take a deterministic stationary policy $\hat{\Phi}$ such that $c_0(x,\hat{\Phi}(x)) - \alpha u^*(x) + \int_S q(dy|x,\hat{\Phi}(x))u^*(y) \le \alpha\varepsilon$ holds for all $x \in S$. Therefore, by (2.15) in Lemma 2.5, $V_0(\hat{\Phi}) \le \int_S \gamma(dy)u^*(y) + \varepsilon$. Consequently, we have $\inf_\Phi V_0(\Phi) \le \int_S \gamma(dy)u^*(y)$, as $\varepsilon > 0$ is arbitrary. On the other hand, Lemma 2.5 implies for any π, $V_0(\pi) \ge \int_S \gamma(dy)u^*(y)$. Therefore, we conclude $\int_S \gamma(dy)u^*(y) = \inf_\pi V_0(\pi) = \inf_\Phi V_0(\Phi)$, as required.

(b) Let us fix arbitrarily $x \in S$ and put $\hat{\gamma}(\Gamma) = I\{x \in \Gamma\}$ for any $\Gamma \in \mathscr{B}(S)$. Clearly $\hat{\gamma}$ satisfies Conditions 2.3 (a). Suppose now there is another solution $v^* \in \mathbf{B}_{w'}(S)$ to the Bellman equation (2.14), then it follows from part (a) that $\inf_\pi V_0(\pi) = u^*(x) = v^*(x)$.

(c) Clearly, u^* is feasible for the linear program (2.16). Suppose now, there is v also feasible for (2.16) and such that $\int_S \gamma(dy)v(y) > \int_S \gamma(dy)u^*(y)$. Then there exist $x \in S$ and $\delta > 0$ so that $v(x) > u^*(x) + \delta$. According to Lemma 2.3 (b), for any π, $v(x) \le V_0(x,\pi)$. Hence, we conclude $u^*(x) < V_0(x,\pi) - \delta$, which contradicts part (a) of this theorem, as desired.

(d) Since u^* solves the DLP (see part (c)), the optimal value of DLP is $\int_S u^*(y)\gamma(dy)$. Therefore, if some v feasible for the DLP is such that $u^*(x) = v(x)$ $\gamma - a.s.$, it solves the DLP as well. Hence we conclude the sufficiency part of the statement.

On the other hand, suppose another solution to the DLP, v, is not equal to u^* $\gamma - a.s.$. Then there are only three possibilities.

Case 1: There exists a measurable $\Gamma \subseteq S$ such that $\gamma(\Gamma) > 0$, $v(x) > u(x)$ on Γ and $v(x) = u(x)$ $\gamma - a.s.$ on $S \setminus \Gamma$. Then $\int_S v(x)\gamma(dx) > \int_S u^*(x)\gamma(dx)$, which contradicts the fact that $\int_S u^*(x)\gamma(dx)$ is the optimal value of the DLP.

Case 2: There exists a measurable $\Gamma \subseteq S$ such that $\gamma(\Gamma) > 0$, $v(x) < u(x)$ on Γ and $v(x) = u(x)$ $\gamma - a.s.$ on $S \setminus \Gamma$. Then $\int_S v(x)\gamma(dx) < \int_S u^*(x)\gamma(x)$, which is a contradiction against the fact that v solves the DLP.

Case 3: There exist measurable $\Gamma_1 \subseteq S$, $\Gamma_2 \subseteq S$ such that $\Gamma_1 \bigcap \Gamma_2 = \emptyset$, $\gamma(\Gamma_1) > 0$, $\gamma(\Gamma_2) > 0$, and $v(x) > u^*(x)$ on Γ_1 and $v(x) < u^*(x)$ on Γ_2. Now let us define another function

$$\hat{v}(x) = I\{x \in S \setminus \Gamma_2\}v(x) + I\{x \in \Gamma_2\}u^*(x),$$

which is obviously feasible for the DLP. Indeed, firstly, it is evident that $\hat{v} \in \mathbf{B}_{w'}(S)$; and secondly for $x \in S \setminus \Gamma_2$,

$$\begin{aligned}
& \frac{1}{\alpha}c_0(x,a) - \hat{v}(x) + \frac{1}{\alpha}\int_S \hat{v}(y)q(dy|x,a) \\
= \ & \frac{1}{\alpha}c_0(x,a) - v(x) + \frac{1}{\alpha}\int_{S\setminus\Gamma_2} v(y)q(dy|x,a) + \frac{1}{\alpha}\int_{\Gamma_2} u^*(y)q(dy|x,a) \\
\geq \ & \frac{1}{\alpha}c_0(x,a) - v(x) + \frac{1}{\alpha}\int_{S\setminus\Gamma_2} v(y)q(dy|x,a) + \frac{1}{\alpha}\int_{\Gamma_2} v(y)q(dy|x,a) \geq 0;
\end{aligned}$$

and for $x \in \Gamma_2$,

$$\begin{aligned}
& \frac{1}{\alpha}c_0(x,a) - \hat{v}(x) + \frac{1}{\alpha}\int_S \hat{v}(y)q(dy|x,a) \\
= \ & \frac{1}{\alpha}c_0(x,a) - u^*(x) + \frac{1}{\alpha}\int_{S\setminus\Gamma_2} v(y)q(dy|x,a) + \frac{1}{\alpha}\int_{\Gamma_2} u^*(y)q(dy|x,a) \\
\geq \ & \frac{1}{\alpha}c_0(x,a) - u^*(x) + \frac{1}{\alpha}\int_{S\setminus\Gamma_2} u^*(y)q(dy|x,a) + \frac{1}{\alpha}\int_{\Gamma_2} u^*(y)q(dy|x,a) \geq 0.
\end{aligned}$$

Then $\int_S \hat{v}(y)\gamma(dy) = \int_{S\setminus\Gamma_2} v(x)\gamma(dx) + \int_{S\setminus\Gamma_2} u^*(x)\gamma(dx) > \int_S v(x)\gamma(dx)$, which is a contradiction against that v is optimal. Hence, the necessity follows. ■

Proof of Lemma 2.6. For simplicity, in the proof of this statement, we shall use the denotations $q(dy|x,\pi) \triangleq \int_A \pi(da|x)q(dy|x,a)$ and $c_0(x,\pi) \triangleq \int_A \pi(da|x)c_0(x,a)$. Given the stationary policy π and the conditions of this statement, according to (Gikhman and Skorokhod 1996, Thm.4, p.364), Kolmogorov's backward equation (2.17) holds. Now we have

$$
\begin{aligned}
& c_0(x,\pi)+\int_S\int_A \pi(da|x)q(dy|x,a)V(y,\pi) \\
= \; & c_0(x,\pi)+\int_S q(dy|x,\pi)\int_0^\infty e^{-\alpha t}\int_S P_y^\pi(\xi_t\in dz)c_0(z,\pi)dt \\
= \; & c_0(x,\pi)+\int_S\int_0^\infty e^{-\alpha t}\int_S q(dy|x,\pi)P_y^\pi(\xi_t\in dz)c_0(z,\pi)dt.
\end{aligned}
$$

According to (2.17), integrating by parts leads to

$$
\begin{aligned}
& \int_0^\infty e^{-\alpha t}\int_S\int_S q(dy|x,\pi)P_y^\pi(\xi_t\in dz)c_0(z,\pi)dt \\
= \; & \int_0^\infty e^{-\alpha t}\frac{d}{dt}\left[\int_S P_x^\pi(\xi_t\in dz)c_0(z,\pi)\right]dt \\
= \; & -E_x^\pi[c_0(\xi_0,\pi)]+\int_0^\infty \alpha e^{-\alpha t}E_x^\pi[c_0(\xi_t,\pi)]dt,
\end{aligned}
$$

and thus the RHS of (2.18) equals the LHS. The interchange of integrals and expectations is legal, as can be easily verified under the conditions of the statement. ■

Proof of Theorem 2.5. (a) In this proof, all the encountered changes of orders of expectations and integrals are legal, as can be easily verified. Let us firstly prove that η^π satisfies the first relation. (Hernández-Lerma and Lasserre 1996, C.10 Prop.) frees us to write down from the Kolmogorov's forward equation (2.10)

$$
\begin{aligned}
P_\gamma^\pi(\xi_t\in\Gamma_S) &= \int_S P_x^\pi(\xi_t\in\Gamma_S)\gamma(dx) \\
&= \int_S \gamma(dx)\left\{I\{x\in\Gamma_S\}+E_x^\pi\left[\int_0^t\int_A \pi(da|\omega,u)q(\Gamma_S\setminus\{\xi_u\}|\xi_u,a)du\right]\right. \\
&\quad \left. -E_x^\pi\left[\int_0^t\int_A \pi(da|\omega,u)q_{\xi_u}(a)I\{\xi_u\in\Gamma_S\}du\right]\right\} \\
&= \gamma(\Gamma_S)+E_\gamma^\pi\left[\int_0^t\int_A \pi(da|\omega,u)q(\Gamma_S\setminus\{\xi_u\}|\xi_u,a)du\right] \\
&\quad -E_\gamma^\pi\left[\int_0^t\int_A \pi(da|\omega,u)q_{\xi_u}(a)I\{\xi_u\in\Gamma_S\}du\right],
\end{aligned}
$$

from which we can write

$$
\begin{aligned}
\eta^\pi(\Gamma_S\times A) &= \alpha\int_0^\infty e^{-\alpha t}P_\gamma^\pi(\xi_t\in\Gamma_S)dt \\
&= \alpha\int_0^\infty e^{-\alpha t}\gamma(\Gamma_S)dt+\alpha\int_0^\infty\left\{e^{-\alpha t}\int_0^t E_\gamma^\pi\left[\int_A\pi(da|\omega,u)\right.\right. \\
&\quad \left.\left.\times q(\Gamma_S\setminus\{\xi_u\}|\xi_u,u)\right]du\right\}dt \\
&\quad -\alpha\int_0^\infty\left\{e^{-\alpha t}\int_0^t E_\gamma^\pi\left[\int_A\pi(da|\omega,u)q_{\xi_u}(a)\right.\right. \\
&\quad \left.\left.\times I\{\xi_u\in\Gamma_S\}\right]du\right\}dt. \qquad (2.44)
\end{aligned}
$$

By integration by parts, for the second term in the last expression, we have

$$
\begin{aligned}
& \alpha \int_0^\infty \left\{ e^{-\alpha t} \int_0^t E_\gamma^\pi \left[\int_A \pi(da|\omega,u) q(\Gamma_S \setminus \{\xi_u\}|\xi_u,u) \right] du \right\} dt \\
= & \left\{ -e^{-\alpha t} \int_0^t E_\gamma^\pi \left[\int_A \pi(da|\omega,u) q(\Gamma_S \setminus \{\xi_u\}|\xi_u,a) \right] du \right\} \Big|_0^\infty \\
& + \int_0^\infty \left\{ e^{-\alpha t} E_\gamma^\pi \left[\int_A \pi(da|\omega,t) q(\Gamma_S \setminus \{\xi_t\}|\xi_t,a) \right] \right\} dt \\
= & \int_0^\infty \left\{ e^{-\alpha t} E_\gamma^\pi \left[\int_A \pi(da|\omega,t) q(\Gamma_S \setminus \{\xi_t\}|\xi_t,a) \right] \right\} dt,
\end{aligned}
$$

because $\{-e^{-\alpha t} \int_0^t E_\gamma^\pi [\int_A \pi(da|\omega,u) q(\Gamma_S \setminus \{\xi_u\}|\xi_u,a)] du\}|_0^\infty = 0$. Indeed, by Conditions 2.2,

$$
E_\gamma^\pi \left[\int_A \pi(da|\omega,u) q(\Gamma_S \setminus \xi_u|\{\xi_u\},a) \right] \leq L E_\gamma^\pi [w(\xi_u)] \leq L \int_S \gamma(dx) h(0,x,u),
$$

with h given by (2.6), which, as can be easily verified, leads to that

$$
\lim_{t \to \infty} (-e^{-\alpha t}) \int_0^t E_\gamma^\pi \left[\int_A \pi(da|\omega,u) q(\Gamma_S \setminus \{\xi_u\}|\xi_u,a) \right] du = 0.
$$

After treating the third term in the last line of (2.44) similarly, following (2.44), we then obtain

$$
\begin{aligned}
& \eta^\pi(\Gamma_S \times A) \\
= & \gamma(\Gamma_S) + \int_0^\infty e^{-\alpha t} E_\gamma^\pi \left[\int_A \pi(da|\omega,t) q(\Gamma_S \setminus \{\xi_t\}|\xi_t,a) \right] dt \\
& - \int_0^\infty e^{-\alpha t} E_\gamma^\pi \left[\int_A \pi(da|\omega,t) q_{\xi_t}(a) I\{\xi_t \in \Gamma_S\} \right] dt \\
= & \gamma(\Gamma_S) + \int_A \int_0^\infty e^{-\alpha t} E_\gamma^\pi [\pi(da|\omega,t) q(\Gamma_S|\xi_t,a)] dt \\
= & \gamma(\Gamma_S) + \int_A \int_0^\infty e^{-\alpha t} \int_S E_\gamma^\pi [\pi(da|\omega,t) q(\Gamma_S|\xi_t,a)|\xi_t = y] P_\gamma^\pi(\xi_t \in dy) dt \\
= & \gamma(\Gamma_S) + \int_A \int_0^\infty e^{-\alpha t} \int_S q(\Gamma_S|y,a) E_\gamma^\pi [\pi(da|\omega,t)|\xi_t = y] P_\gamma^\pi(\xi_t \in dy) dt \\
= & \gamma(\Gamma_S) + \int_{S \times A} \int_0^\infty e^{-\alpha t} q(\Gamma_S|y,a) E_\gamma^\pi [\pi(da|\omega,t) I\{\xi_t \in dy\}] dt \\
= & \gamma(\Gamma_S) + \frac{1}{\alpha} \int_{S \times A} q(\Gamma_S|y,a) \eta^\pi(dy \times da),
\end{aligned}
$$

where the third equality follows from the definition of conditional expectation, and the second last equality follows from that for any fixed $\Gamma_A \in \mathscr{B}(A)$, $E_\gamma^\pi[\pi(\Gamma_A|\omega,t)|\xi_t = y]$ is the Radon-Nykodym's derivative of $E_\gamma^\pi[\pi(\Gamma_A|\omega,t) I\{\xi_t \in dy\}]$, a measure on S, with respect to $P_\gamma^\pi(\xi_t \in dy)$. In more details, while the absolute continuity part is trivial, by the very definition of conditional expectation, for any $\Gamma_S \in \mathscr{B}(S)$, we have

$$\begin{aligned}
& E_\gamma^\pi[\pi(\Gamma_A|\omega,t)I\{\xi_t \in \Gamma_S\}] \\
= & \int_S E_\gamma^\pi[\pi(\Gamma_A|\omega,t)I\{\xi_t \in \Gamma_S\}|\xi_t = y]P_\gamma^\pi(\xi_t \in dy) \\
= & \int_S I\{y \in \Gamma_S\}E_\gamma^\pi[\pi(\Gamma_A|\omega,t)|\xi_t = y]P_\gamma^\pi(\xi_t \in dy) \\
= & \int_{\Gamma_S} E_\gamma^\pi[\pi(\Gamma_A|\omega,t)|\xi_t = y]P_\gamma^\pi(\xi_t \in dy).
\end{aligned}$$

Now let us show η^π satisfies the second relation. We have

$$\begin{aligned}
\int_S w(x)\eta^\pi(dx \times A) &= \alpha\int_0^\infty e^{-\alpha t}E_\gamma^\pi[w(\xi_t)]dt \\
&\leq \alpha\int_0^\infty e^{-\alpha t}\int_S h(0,x,t)\gamma(dx) \\
&= \frac{\alpha\int_S \gamma(dx)w(x) + b}{\alpha - \rho} < \infty,
\end{aligned}$$

where the first inequality follows from Theorem 2.1, with h given by (2.6).

(b) Recall that for any $\Gamma_S \in \mathscr{B}(S)$ and $\Gamma_A \in \mathscr{B}(A)$

$$\chi_t^\pi(\Gamma_S \times \Gamma_A) = \int_{\Gamma_S} E_\gamma^\pi[\pi(\Gamma_A|\omega,t)|\xi_t = y]P_\gamma^\pi(\xi_t \in dy)$$

from which we can take $E_\gamma^\pi[\pi(\Gamma_A|\omega,t)|\xi_t = y]$ as the stochastic kernel on A given $y \in S$ so that it, together with $P_\gamma^\pi(\xi_t \in dy)$, determines the probability measure $\chi_t^\pi(dx \times da)$. Now, one can observe easily that under the conditions of the statement, $P_\gamma^\pi(\xi_t \in S) = 1$ (see Theorem 2.1) and $E_\gamma^\pi[\pi(A(y)|\omega,t)|\xi_t = y] = 1$. Therefore, according to (Dynkin and Yushkevich 1979, p.88, Thm.1), we obtain that $\chi_t^\pi(K) = 1$, which implies $\eta^\pi(K) = 1$, as required.

(c) As was mentioned in the statement, according to (Hernández-Lerma and Lasserre 1996, D.8 Prop.), π, a stochastic kernel on A given $y \in S$, which satisfies (2.22), is of existence. Due to (Hernández-Lerma and Lasserre 1996, D.8 Prop.), one can take $\pi(da|y)$ concentrated on $A(y)$. Then evidently, π is a stationary policy, and one can introduce its occupation measure η^π. Now for the claimed statement, we shall prove that for any measurable bounded function u on $S \times A$,

$$\int_{S\times A} u(x,a)\eta(dx \times da) = \int_{S\times A} u(x,a)\eta^\pi(dx \times da).$$

Define $V(x) \triangleq E_x^\pi[\int_0^\infty e^{-\alpha t}\int_A \pi(da|\xi_t)u(\xi_t,a)dt]$. Then, by Lemma 2.6, $V(x)$ satisfies the following equation:

$$\alpha V(x) = \int_A u(x,a)\pi(da|x) + \int_S \int_A \pi(da|x) q(dy|x,a) V(y). \tag{2.45}$$

Then one can easily verify that the conditions imposed in the statement imply

$$\int_S \eta(dx \times A) \left\{ \int_A \int_S q(dy \setminus \{x\}|x,a)\pi(da|x)|V(y)| + \int_A q_x(a)\pi(da|x)|V(x)| \right\} < \infty,$$

which legalizes the interchange of integrals in the forthcoming calculations.

Due to (2.20), (2.22) and (2.45), we have

$$\begin{aligned}
& \int_{S\times A} u(y,a)\eta(dy \times da) \\
= & \int_S \eta(dy \times A) \int_A u(y,a)\pi(da|y) \\
= & \int_S \eta(dy \times A) \left\{ \alpha V(y) - \int_S \int_A q(dz|y,a)\pi(da|y)V(z) \right\} \\
= & \int_S \eta(dy \times A)\alpha V(y) - \int_S \eta(dy \times A) \int_S \int_A q(dz|y,a)\pi(da|y)V(z)
\end{aligned}$$

$$\begin{aligned}
= & \int_S \left\{ \gamma(dy) + \frac{1}{\alpha} \int_{S\times A} q(dy|z,a)\eta(dz \times da) \right\} \alpha V(y) \\
& - \int_S \eta(dy \times A) \int_S \int_A q(dz|y,a)\pi(da|y)V(z) \\
= & \int_S \gamma(dy)\alpha V(y) + \int_S \int_{S\times A} q(dy|z,a)\eta(dz \times da)V(y) \\
& - \int_S \eta(dy \times A) \int_S \int_A q(dz|y,a)\pi(da|y)V(z) \\
= & \alpha \int_S \gamma(dy)V(y).
\end{aligned}$$

On the other hand, we have

$$\alpha \int_S \gamma(dy)V(y) = \int_{S\times A} u(y,a)\eta^{\pi}(dx \times da).$$

Indeed, with in mind the definition of χ_t^π, one can write

$$\begin{aligned}
\int_{S\times A} u(y,a)\eta^\pi(dx\times da) &= \alpha\int_{S\times A} u(y,a)\int_0^\infty e^{-\alpha t}\chi_t^\pi(dx\times da)dt \\
&= \alpha\int_0^\infty e^{-\alpha t}\int_S \gamma(dx)\int_S P_x^\pi(\xi_t\in dy) \\
&\quad \int_A u(y,a)E_x^\pi\left[\pi(da|\xi_t)|\xi_t=y\right]dt \\
&= \alpha\int_0^\infty e^{-\alpha t}\int_S \gamma(dx)E_x^\pi\left[\int_A u(\xi_t,a)\pi(da|\xi_t)\right]dt \\
&= \alpha\int_S \gamma(dx)V(x).
\end{aligned}$$

Therefore, we finally can conclude that for any bounded measurable u on $S\times A$,

$$\int_{S\times A} u(x,a)\eta(dx\times da) = \int_{S\times A} u(x,a)\eta^\pi(dx\times da).$$

Now it only remains to put $u(x,a)$ as an indicator function to induce the statement.

(d) If $\hat{\pi}$ is not a version of π, then there exists $\Gamma_S\in\mathscr{B}(S)$ with $\eta(\Gamma_S\times A)>0$ and $\Gamma_A\in\mathscr{B}(A)$ such that $\hat{\pi}(\Gamma_A|x)\neq\pi(\Gamma_A|x)$ on Γ_S. Clearly, there exist disjoint partitions of Γ_S, namely, Γ_S^1 and Γ_S^2, such that $\hat{\pi}(\Gamma_A|x)>\pi(\Gamma_A|x)$ on Γ_S^1, and $\hat{\pi}(\Gamma_A|x)<\pi(\Gamma_A|x)$ on Γ_S^2, and at least one of them has a positive measure with respect to $\tilde{\eta}(dx)\triangleq\eta(dx\times A)$. Without loss of generality, as the reasoning will hold with obvious modifications in the opposite case, let us take $\eta(\Gamma_S^1\times A)>0$. If $\eta^{\hat{\pi}}(dx\times A)\neq\eta(dx\times A)$, then we have a contradiction against the fact $\eta^{\hat{\pi}}=\eta$. Therefore, we shall suppose $\eta^{\hat{\pi}}(dx\times A)=\eta(dx\times A)$ identically. But then

$$\begin{aligned}
\eta^{\hat{\pi}}(\Gamma_S^1\times\Gamma_A) &= \alpha E_\gamma^{\hat{\pi}}\left[\int_0^\infty e^{-\alpha t}I\{\xi_t\in\Gamma_S^1\}\hat{\pi}(\Gamma_A|\xi_t)dt\right] = \int_{\Gamma_S^1}\eta^{\hat{\pi}}(dx\times A)\hat{\pi}(\Gamma_A|x) \\
&= \int_{\Gamma_S^1}\eta(dx\times A)\hat{\pi}(\Gamma_A|x) > \int_{\Gamma_S^1}\eta(dx\times A)\pi(\Gamma_A|x) = \eta(\Gamma_S^1\times\Gamma_A),
\end{aligned}$$

where the last equality follows from (2.22). However, this yields a contradiction against $\eta^{\hat{\pi}}=\eta$. Therefore, one can conclude the statement.

(e) Clearly, due to part (a), $\eta^\pi(dx\times da)$ satisfies (2.20) and (2.21). Therefore, equation (2.23) is solvable with a solution $\tilde{\eta}^\pi(dx)\triangleq\eta^\pi(dx\times A)$ subject to (2.24). Now suppose that $\hat{\eta}(dx)$ is another solution to equation (2.23) subject to (2.24) and different from $\tilde{\eta}^\pi(dx)$. Then let us define $\eta(dx\times da)$ by $\eta(\Gamma_S\times\Gamma_A)\triangleq\int_{\Gamma_S}\hat{\eta}(dx)\pi(\Gamma_A|x)$, where $\Gamma_S\in\mathscr{B}(S)$ and $\Gamma_A\in\mathscr{B}(A)$. One can easily check that (2.20) and (2.21) are satisfied by η. Therefore, applying part (c), there exists a stationary policy $\hat{\pi}$ such that $\eta=\eta^{\hat{\pi}}$.

Consequently, the following relation holds:

$$\eta(\Gamma_S \times \Gamma_A) = \int_{\Gamma_S} \eta^{\tilde{\pi}}(dx \times A)\tilde{\pi}(\Gamma_A|x) = \int_{\Gamma_S} \hat{\eta}(dx)\pi(\Gamma_A|x), \tag{2.46}$$

where $\Gamma_S \in \mathscr{B}(S)$, $\Gamma_A \in \mathscr{B}(A)$, and the last equality follows from the definition of η. Now putting $\Gamma_A = A$ in (2.46) yields that for any $\Gamma_S \in \mathscr{B}(S)$,

$$\hat{\eta}^{\tilde{\pi}}(\Gamma_S) \triangleq \eta^{\tilde{\pi}}(\Gamma_S \times A) = \eta(\Gamma_S \times A) = \hat{\eta}(\Gamma_S),$$

by putting which back into the second expression of (2.46), we obtain

$$\int_{\Gamma_S} \hat{\eta}(dx)\tilde{\pi}(\Gamma_A|x) = \int_{\Gamma_S} \hat{\eta}(dx)\pi(\Gamma_A|x),$$

from which we conclude that for any $\Gamma_A \in \mathscr{B}(A)$, $\pi(\Gamma_A|x) = \tilde{\pi}(\Gamma_A|x)$ a.s. with respect to $\hat{\eta}^{\tilde{\pi}}(dx) \triangleq \eta^{\tilde{\pi}}(dx \times A)$. With this relation in mind, in addition to the definition of occupation measure, as well as (2.4), one can easily show that $P^{\pi}_{\gamma} = P^{\tilde{\pi}}_{\gamma}$, implying that $\hat{\eta}^{\tilde{\pi}}(dx) = \hat{\eta}^{\pi}(dx)$, due to the very definition of occupation measure. On the other hand, we also have $\hat{\eta}^{\tilde{\pi}}(dx) = \hat{\eta}(dx)$. Therefore, one can finally conclude that $\hat{\eta}(dx) = \hat{\eta}^{\pi}(dx)$, a desired contradiction. ■

Proof of Lemma 2.7. The "only if" part: Let $M_n \xrightarrow{f} M$ be given as in the statement, where $M_n, M \in \mathscr{M}_f(K)$. Then $\lim_{n\to\infty} \int_S f(x)M_n(dx \times A) = \int_S f(x)M(dx \times A)$. Now let us take an arbitrary bounded continuous $g(x,a)$ on K. Then

$$\begin{aligned}
\lim_{n\to\infty} \int_K g(x,a)\tilde{M}_n(dx \times da) &= \lim_{n\to\infty} \int_K g(x,a)\frac{f(x)M_n(dx \times da)}{\int_S f(x)M_n(dx \times A)} \\
&= \frac{\lim_{n\to\infty} \int_K g(x,a)f(x)M_n(dx \times da)}{\lim_{n\to\infty} \int_S f(x)M_n(dx \times A)} \\
&= \frac{\int_K g(x,a)f(x)M(dx \times da)}{\int_S f(x)M(dx \times A)} \\
&= \int_K g(x,a)\tilde{M}(dx \times da),
\end{aligned}$$

meaning that $\tilde{M}_n \to \tilde{M}$.

The "if" part: Let $\tilde{M}_n \to \tilde{M}$, where $\tilde{M}, \tilde{M}_n, \in \mathscr{M}(K)$. Then $\lim_{n\to\infty} \int_K \tilde{M}_n(dx \times da)\frac{1}{f(x)} = \int_S \tilde{M}_0(dx \times A)\frac{1}{f(x)}$, simply because $\frac{1}{f(x)}$ is bounded and continuous on S. Now let us take an arbitrary continuous function $u(x,a)$ on K such that $\sup_{x\in S} \frac{\sup_{a\subset A(x)} |u(x,a)|}{f(x)} < \infty$. Then the function $g(x,a) \triangleq \frac{u(x,a)}{f(x)}$ is clearly bounded and continuous on K. Therefore,

$$\lim_{n\to\infty}\int_K u(x,a)M_n(dx\times da) = \lim_{n\to\infty}\frac{\int_K u(x,a)\frac{\tilde{M}_n(dx\times da)}{f(x)}}{\int_S \frac{\tilde{M}_n(dx\times A)}{f(x)}} = \frac{\int_K \frac{u(x,a)}{f(x)}\tilde{M}(dx\times da)}{\int_S \frac{1}{f(x)}\tilde{M}(dx\times A)}$$
$$= \int_K u(x,a)M(dx\times da),$$

where M and $M_n, n=1,2,\dots$ are transformed via (2.26). Consequently, $M_n \xrightarrow{f} M$, as required. ■

Proof of Theorem 2.6. Let us consider a sequence of $\eta_n \in \mathscr{D}$ such that $\eta_n \xrightarrow{w} \eta$, where $\eta_n, \eta \in \mathscr{M}_w(K)$, and show that $\eta \in \mathscr{D}$. This will suffice our claim, because of Remark 2.10. To this end, due to Remark 2.9, we only need verify that η satisfies (2.20) and (2.21). Clearly, (2.21) holds for η, simply because it holds for η_n, together with the finite upper bound in (2.21). Therefore, it remains to check the validity of (2.20). Define $\tilde{\eta}(dx) \triangleq \gamma(dx) + \frac{1}{\alpha}\int_K q(dx|y,a)\eta(dy\times da)$ a measure on S. Then for any continuous and bounded $u(x)$, we have

$$\lim_{n\to\infty}\int_S u(x)\eta_n(dx\times A)$$
$$= \lim_{n\to\infty}\left\{\int_S u(x)\gamma(dx) + \frac{1}{\alpha}\int_{S\times A}\int_S u(x)q(dx|y,a)\eta_n(dy\times da)\right\}$$
$$= \int_S u(x)\gamma(dx) + \frac{1}{\alpha}\int_S u(x)\int_{S\times A} q(dx|y,a)\eta(dy\times da)$$
$$= \int_S u(x)\tilde{\eta}(dx),$$

where the second equality holds due to the following: $\int_S u(x)q(dx|y,a)$ is continuous in (y,a), $\sup_{x\in S}\frac{\sup_{a\subset A(x)}|\int_S u(x)q(dx|y,a)|}{w(y)} < \infty$ (because u is bounded and $\bar{q}_x$ is w-bounded, see Conditions 2.2), and $\eta_n \xrightarrow{w} \eta$. Hence, we obtain that $\eta_n(dx\times A) \to \tilde{\eta}(dx)$, where the convergence is in the usual weak topology. On the other hand, vividly, we also have $\eta_n(dx\times A) \xrightarrow{w} \eta(dx\times A)$, which follows from $\eta_n(dx\times da) \xrightarrow{w} \eta(dx\times da)$ and that any $f(x)$ continuous on S is automatically continuous on K. This further leads to that $\eta_n(dx\times A) \to \eta(dx\times A)$, because the usual weak convergence is weaker than the w-weak convergence (recall $w\geq 1$). Now, one can infer from the uniqueness of the weak limit that $\tilde{\eta}(dx) = \eta(dx\times A)$. In other words, for any fixed $\Gamma_S \in \mathscr{B}(S)$, $\eta(\Gamma_S\times A) = \gamma(\Gamma_S) + \frac{1}{\alpha}\int_K q(\Gamma_S|y,a)\eta(dy\times da)$, meaning that (2.20) is satisfied by η. ■

Proof of Theorem 2.7. Note that Conditions 2.10 are stronger than Conditions 2.9. Therefore, under the imposed conditions, we have that $\mathscr{D}$ is w-weakly closed and convex. Then according to (Rockafellar 1974, Thm.17(a)) as well as Example 1", p.45

therein, whose validity requires further Slater's condition, we have

$$\begin{aligned}\inf(PLP\ (2.27)) &= \inf(PLP\ (2.28)) \\ &= \sup_{g_n \geq 0, n=1,\dots,N} \inf_{\eta \in \mathscr{D}} \left\{ \frac{1}{\alpha} \int_K \bar{c}(x,a)\eta(dx \times da) \right\}, \quad (2.47)\end{aligned}$$

where $\bar{c}(x,a) \triangleq c_0(x,a) + \sum_{n=1}^N g_n c_n(x,a) - \sum_{n=1}^N g_n d_n \alpha$, and the term inside the parenthesis is the Lagrangean function, see (Rockafellar 1974, (4.2) in p.18, (5.1) in p.23).

Now for any fixed $g_n \geq 0$, $n = 1, \dots, N$, the problem of

$$\frac{1}{\alpha} \int_K \bar{c}(x,a)\eta(dx \times da) \to \min_{\eta \in \mathscr{D}} \quad (2.48)$$

takes the same form as PLP (2.27): it is simply the case of $N = 0$ (compare with PLP 2.28)). Its DLP takes the form of DLP (2.30) (see also DLP (2.16)):

$$\int_S \gamma(dx)v(x) \to \max_{v \in \mathscr{V}_0} \quad (2.49)$$

$$s.t. \quad \frac{1}{\alpha}\bar{c}(x,a) - v(x) + \frac{1}{\alpha} \int_S v(y)q(dy|x,a) \geq 0. \quad (2.50)$$

By Remark 2.7, there is a solution $\bar{u}^* \in \mathbf{B}_{w'}(S)$ (as well as $\bar{u}^* \in \mathbf{B}_w(S)$) to the following Bellman equation

$$\alpha \bar{u}^*(x) = \inf_{a \in A(x)} \left\{ \bar{c}(x,a) + \int_S \bar{u}^*(y)q(dy|x,a) \right\}. \quad (2.51)$$

Now by Theorem 2.4, we have

$$\begin{aligned}\inf(PLP\ (2.48)) &= \int_S \gamma(dx)\bar{u}^*(x) = \sup(DLP\ (2.49)) \\ &= \sup_{v \in \mathscr{V}_0} \int_S \gamma(dx)v(x) \\ &\quad \text{(s.t. constraints in (2.30) with fixed } g_n \geq 0, n = 1, \dots, N),\end{aligned}$$

taking sup with respect to $g_n \geq 0, n = 1, \dots, N$ to the both sides of which will not break the equality, leading to

$$\sup_{g_n \geq 0, n=1,\dots,N} \inf(PLP\ (2.48)) = \sup(DLP\ (2.30)).$$

On the other hand, (2.47) asserts that $\inf(PLP\ (2.27)) = \sup_{g_n \geq 0, n=1,\dots,N} \inf(PLP\ (2.48))$. Therefore, we obtain $\sup(DLP\ (2.30)) = \inf(PLP\ (2.27))$, as required. ■

Proof of Lemma 2.8. Let us firstly prove an auxiliary statement, asserting that $\mathscr{D}$ is

w'-weakly precompact in $\mathcal{M}_{w'}(K)$. By the way, the fact of $\mathcal{D}$ being a subset of $\mathcal{M}_{w'}(K)$ follows from the imposed conditions such as Conditions 2.6 (a,b). Due to Remark 2.10, to this end, it suffices to show its image, $\tilde{\mathcal{D}} = Q_{w'}(\mathcal{D})$, whose elements are denoted as $\tilde{\eta}$ in this proof, to be precompact in the usual weak topology. Conditions 2.11 (b) imply that $\frac{w}{w'}$ is strictly unbounded, because of (Hernández-Lerma and Lasserre 1996, Rem.5.7.5(b)). Now we have

$$\begin{aligned}
\sup_{\tilde{\eta}} \int_S \frac{w(x)}{w'(x)} \tilde{\eta}(dx \times A) &= \sup_{\tilde{\eta}} \int_S w(x) \frac{\eta(dx \times A)}{\int_S w'(x)\eta(dx \times A)} \\
&\leq \frac{\alpha \int_S w(x)\gamma(dx) + b}{\alpha - \rho} < \infty,
\end{aligned}$$

where we recall that $w'(x) \geq 1$ (see Conditions 2.6). From this we can infer from (Hernández-Lerma and Lasserre 1996, E.8 Prop.) that $\tilde{\mathcal{D}}$ is tight. This, according to Prokhorov's theorem (Hernández-Lerma and Lasserre 1996, E.8 Thm.), further leads to that $\tilde{\mathcal{D}}$ is precompact in the usual weak topology. Therefore, the auxiliary statement is proved.

Now due to the auxiliary statement, to show the claimed statement, it suffices to show that $\mathcal{D}$ is w'-weakly closed. However, under the imposed conditions, this can be proved in exactly the same manner as for Theorem 2.6, and hence its proof is omitted here. ■

Proof of Theorem 2.8. As noted above, the elements of $\mathcal{D}$ also satisfy (2.21) with w, ρ, and b being replaced by w', ρ' and b', respectively, as can be checked easily. Now the non-empty subset $\mathcal{D}^{w'} \subseteq \mathcal{D}$ of occupation measures, whose elements η are feasible and satisfy $\int_K w'(x)\eta(dx \times da) < \infty$, has the form

$$\mathcal{D}^{w'} = \left\{ \eta \in \mathcal{D} : \int_K c_n(x,a)\eta(dx \times da) \leq d_n, n = 1, \ldots, N \right\},$$

which is w'-weakly closed in $\mathcal{D}$. Indeed, to see why, let us suppose $\eta_j \xrightarrow{w'} \eta$, where $\eta_j \in \mathcal{D}^{w'}$, and show $\int_K c_n(x,a)\eta(dx \times da) \leq d_n$, $n = 1, \ldots, N$. Let us fix $n \in \{1, 2, \ldots, N\}$. Since $c_n(x,a)$ are lower semicontinuous on K (Conditions 2.11), there exists a nondecreasing sequence of continuous functions on K, namely, $c_n^m(x,a)$, such that $c_n^m(x,a) \uparrow$

$c_n(x,a)$, see for example, (Piunovskiy 1997, Thm. A1.14). Then,

$$\begin{aligned}
\int_K \eta(dx\times da)c_n(x,a) &= \int_K \eta(dx\times da)\lim_{m\to\infty} c_n^m(x,a) \\
&= \lim_{m\to\infty}\int_K \eta(dx\times da)c_n^m(x,a) \\
&\quad \text{(Lebesgue dominated convergence theorem)} \\
&= \lim_{m\to\infty}\lim_{j\to\infty}\int_K \eta_j(dx\times da)c_n^m(x,a) \\
&\quad (c_n^m \text{ is continuous and } w'\text{-bounded}) \\
&\leq d_n,
\end{aligned}$$

which implies that $\mathscr{D}^{w'}$ is w'-weakly closed. However, according to Lemma 2.8, $\mathscr{D}$ is also w'-weakly compact. Therefore, $\mathscr{D}^{w'}$ is compact. Similarly, we can show that $\forall\, r\in\mathbb{R}$, $\left\{\eta\in\mathscr{D}^{w'}:\int_K c_0(x,a)\eta(dx\times da)\leq r\right\}$ is closed, from which one can infer that the function $\eta\to\int_K c_0(x,a)\eta(dx\times da)$ is lower semicontinuous in the w'-weak topology. Therefore, problem $\int_K c_0(x,a)\eta(dx\times da)\to\min_{\eta\in\mathscr{D}^{w'}}$, equivalent to problem (2.13), is solvable, see (Aliprantis and Border 2007, Thm.2.43). Suppose the minimizer is η^*. Now, according to Theorem 2.5, one can take an associated (possibly randomized) stationary policy π^*, which is constrained-optimal, as required. ■

Chapter 3

Fluid approximation: Birth-and-Death processes

3.1 Introduction

An important class of the general CTMDPs like those considered in Chapter 2 are CTMDPs with local transitions, which we study in this chapter. Markovian models with local transitions such as Birth-and-Death processes play important roles in Queueing theory. While the underlying stochastic models often tend to be difficult for analytical studies, their fluid models might be taken to provide reasonable approximations. In various contexts, such fluid approximations have been shown to provide a powerful tool. For instance, one may deduce the stability of stochastic queueing networks from that of their fluid models, see Chen (1995). On the other hand, in the context of optimal control, problems of stochastic nature can be satisfactorily solved via their fluid models, see Avrachenkov et al (2005, 2010); Bäuerle (2002); Clancy and Piunovskiy (2005); Gajrat and Hordijk (2005); Piunovskiy (2004b); Piunovskiy and Clancy (2008); Verloop (2009), where meaningful examples in epidemics, queueing networks and telecommunications can be found. We emphasize that in this context, fluid models are often used without analytical justifications.

This chapter is concerned with analytical justifications of fluid approximations to controlled Birth-and-Death processes in the context of optimal control, where one is primarily interested in the performance functionals. As mentioned in Chapter 1, in the general sense fluid models could behave quite differently from stochastic models. Therefore, it is natural to justify analytically such fluid approximations by showing the convergence of a sequence of performance functionals of properly scaled stochastic models to that of the fluid model, with respect to the scaling parameter, say n. The description of this fluid scaling will be given in Section 3.2. In principle, one approach for doing so could be based on the trajectory-wise convergence as proved in Chen and Mandelbaum (1991, 1994); Chen (1996); Ethier and Kurtz (1986); Mandelbaum and Pats (1995); Mandelbaum et al (1998); and examples following this approach include Bäuerle (2000); Pang and Day (2007). In general, one weakness of this approach could be its ineffectiveness in dealing with the case of an infinite horizon, because the trajectory-wise convergence is usually established uniformly on a fixed finite horizon, see Section 3.6 below for more discussions. In comparison, another approach, which is based on the investigations of algebraic equations of the dynamic programming type and without referring to any trajectory-wise convergence, is more powerful in dealing with infinite horizon cases, and it also allows to estimate the rate of convergence (performance functional-wise), see Piunovskiy (2009a,b). In fact, to out best knowledge, it seems that the current literature on justifications of fluid approximations does not provide estimates of the rate of convergence, with few exceptional cases like Gajrat et al (2003); Piunovskiy (2009a,b), where Gajrat et al (2003) studied tandem queues in discrete time on a finite horizon, Piunovskiy (2009a) investigated a single-server-multiclass queueing network with constant arrival rates of jobs, and (Piunovskiy 2009b, Sec.5) gently produced results for controlled Birth-and-Death processes with bounded transition rates and cost rate, and the transition rates are also separated from zero. However, we deem that it is of importance and interest to estimate the rate of convergence, as to some extent, it provides the accuracy of fluid approximations. This fact, alongside the rich applications of Birth-and-Death processes, motivates the present research, carried out by following the second approach mentioned above.

The main contributions of this chapter are double-folded. Firstly, we show via an example that when the transition rates of a controlled Birth-and-Death process are not separated from zero, it might happen that the sequence of performance functionals of the scaled stochastic models does not converge to that of the standard fluid model. Secondly, we propose an alternative fluid model (see Section 3.2) after a change of time scale for the standard fluid model, and justify this refined fluid model with an estimate for the rate of convergence in a closed-form expression involving only primary data. This chapter develops the ideas briefly described in (Piunovskiy 2009b, Sec.5) but is based on much weaker conditions, by allowing the cost rate and transition rates to be unbounded and not separated from zero, which we deem important for practical applications. Indeed, as will be seen in Section 3.2, for the concerned problem in this chapter, the refined fluid model is identical to the standard fluid model if the transition rates are separated from zero, meaning that results obtained in this chapter generalize those obtained in (Piunovskiy 2009b, Sec.5).

In greater details, the main result of this chapter looks like follows. Consider a controlled Birth-and-Death process with an absorbing state of zero, for which one aims at minimizing the expected total cost up to the absorption. Suppose the optimal control problem for the refined fluid model is solved, and the corresponding ε-optimal feedback control policy is built. If one applies that control policy (with natural modifications) to the underlying scaled Birth-and-Death process, then the performance functional will be nearly optimal, that is, $[\delta(\varepsilon)+\hat{\varepsilon}(n)]$-optimal, where $\delta(\varepsilon)\to 0$ as $\varepsilon\to 0$ and $\hat{\varepsilon}(n)\to 0$ as the scaling parameter $n\to\infty$. Moreover, we provide the explicit formulae for $\delta(\varepsilon)$ and the upper boundary of $\hat{\varepsilon}(n)$ in terms of n and the parameters of the model.

The rest of this chapter is organized as follows: Section 3.2 contains the description of the stochastic model under study, its standard fluid model, as well as the newly proposed refined fluid model. In Section 3.3, we impose the main conditions required for the results in this chapter, and provide auxiliary statements. Section 3.4 consists of the main results. Section 3.5 provides some illuminating examples, following which we finish this chapter with comments and conclusions in Section 3.6. The proofs of all the statements are included in Section 3.7 at the very end.

3.2 Description of mathematical models and problem statement

Scaled stochastic model: Following the standard practice, with the scaling parameter $n=1,2,\ldots$, the concerned (scaled) controlled Birth-and-Death process $\{{}^nY_u, u\geq 0\}$ is defined by the following elements: ${}^nS\equiv S=\{0,1,\ldots\}$ is the state space, ${}^nA\equiv A$ is the action space (arbitrary Borel), whose element is denoted by a, ${}^nC(i,a)\triangleq c(\frac{i}{n},a)$ is the cost rate, and ${}^nQ=[{}^nq_{i,j}(a)]_{i,j\in S}$ is the following tri-diagonal matrix of transition rates with components

$$\begin{aligned}
&{}^nq(j|i,a)=0 \text{ if } |i-j|>1;\\
&{}^nq(i+1|i,a)=n\lambda(\frac{i}{n},a);\\
&{}^nq(i-1|i,a)=n\mu(\frac{i}{n},a);\\
&{}^nq(i|i,a)=-({}^nq(i+1|i,a)+{}^nq(i-1|i,a)).
\end{aligned}$$

Here, real measurable functions λ, μ and c are defined on $\mathbb{R}^0_+\times A$, and satisfy

$$\mu(0,a)=\lambda(0,a)=c(0,a)\equiv 0$$

and

$$\mu(x,a)>0,\ \lambda(x,a)\geq 0 \text{ on } (\mathbb{R}^0_+\setminus\{0\})\times A.$$

Clearly, state zero is the absorbing state for $\{{}^nY_u, u\geq 0\}$.

Fixing the initial state ${}^nY_0=i\in S$, a control policy Φ, defined as a mapping from S to A, determines a unique probability measure on the space of trajectories (the canonical space), see Chapter 2, (Kitaev and Rykov 1995, Chap.4) or Piunovskiy (1998) for a rigorous construction. Let us denote E_i^Φ the corresponding expectation operator. In fact, one should put ${}^n\Phi$ and ${}^nE_i^{{}^n\Phi}$ to reflect the scaling. However, we omit the extra superscript n, as far as the context makes it clear about the scaling, and this practice applies also to Chapter 4 and Chapter 5. Then the optimal control problem under consideration looks like follows:

$$ {}^nW^\Phi(i) = E_i^\Phi\left[\int_0^\infty {}^nC({}^nY_u, \Phi({}^nY_u))du\right] \to \inf_\Phi. \tag{3.1}$$

Here, we restrict ourselves to the class of deterministic stationary policies, because under rather general conditions, they are sufficient for solving Markovian optimization problems. For example, that is the case if functions λ, μ and g are uniformly bounded and $\inf_{x>0, a\in A}\frac{\mu(x,a)}{\lambda(x,a)} > 1$. Recall also Theorem 2.4 in Chapter 2. When $n = 1$, the stochastic model is unscaled for which we often omit the index $n = 1$.

The above described scaling is called a standard fluid scaling. Its intuitive meaning is best illustrated via the following example. Suppose Y_t represents the amount of data in the router buffer, measured in Gigabit. Accordingly, $\lambda(i,a)$ represents the average amount of data arrived at the router per unit of time. Now increasing $n = 1000$ means we measure data in Megabit. Therefore, for the 1000th stochastic model, ${}^nY_t$ stands for the amount of data measured in Megabit, and accordingly we must have the arrival rate of $n\lambda(\frac{i}{1000}, a)$ Megabits per arrive per unit of time. In other words, as we increase n, the data are measured more and more accurately, and we observe the stochastic model with better and better "amplifiers". Alternatively but equivalently, one can view the fluid scaling as time acceleration and the state aggregation, see the discussion in Mandelbaum and Pats (1995). Another intuitive meaning (in the context of manufacturing) of this standard fluid scaling is in Chen and Mandelbaum (1994).

Standard fluid model: Let $y(\tau)$ stand for the instantaneous state in the following introduced standard fluid model, for which we denote φ, a measurable mapping from $\mathbb{R}_+^0$ to A, a feedback control policy. Then the standard fluid model can be concisely described by the following problem analogous to problem (3.1):

$$\int_0^\infty c(y, \varphi(y))d\tau \to \inf_\varphi \tag{3.2}$$

$$\text{subject to} \quad \frac{dy}{d\tau} = \lambda(y, \varphi(y)) - \mu(y, \varphi(y)), \text{ with a given initial state } y(0).$$

Remark 3.1 *In formula (3.2), $y(\tau)$ is a function of time, but we omit the argument τ for brevity. the same concerns (3.3) and similar expressions.*

We emphasize that it is problem (3.2) that usually got taken for analysis instead of problem (3.1), which often appears less tractable. Then in (Piunovskiy 2009b, Sec.5), basically, for bounded g, and bounded μ and λ that are also separated from zero, this standard fluid model was justified to provide reasonable approximations to the underlying (scaled) stochastic model. However, when μ and λ are not separated from zero, it might happen that the fluid process $\{y(\tau), \tau \geq 0\}$ behaves qualitatively different from the scaled stochastic process $\{^n Y_u, u \geq 0\}$ on the underlying infinite horizon in that $\{y(\tau), \tau \geq 0\}$ might never get absorbed at zero, while the absorption happens to $\{^n Y_u, u \geq 0\}$ *a.s.*, see Example 3.1 below for more details. On such occasions, as will be seen below, this standard fluid model provides very inaccurate approximations, giving rise to the quest for a refined fluid model.

Refined fluid model: The refined fluid model can be summarized as the following problem:

$$\int_0^\infty \frac{c(y, \varphi(y))}{\lambda(y, \varphi(y)) + \mu(y, \varphi(y))} dt \to \inf_\varphi \tag{3.3}$$

$$\text{subject to} \quad \frac{dy}{dt} = \frac{\lambda(y, \varphi(y)) - \mu(y, \varphi(y))}{\lambda(y, \varphi(y)) + \mu(y, \varphi(y))}, \text{ with a given initial state } y(0).$$

For later reference, let us introduce the following denotation:

$$v^\varphi(y_0) \triangleq \int_0^\infty \frac{c(y, \varphi(y))}{\lambda(y, \varphi(y)) + \mu(y, \varphi(y))} dt = \int_0^{y_0} \frac{c(z, \varphi(z))}{\mu(z, \varphi(z)) - \lambda(z, \varphi(z))} dz, \tag{3.4}$$

if the initial state in the refined fluid model is given by $y(0) = y_0$, for which the zero index is usually omitted.

Clearly, problems (3.2) and (3.3) are identical, if $\inf_{a \in A, y>0}[\mu(y,a) + \lambda(y,a)] \geq \delta > 0$; in this case, $v^\varphi(y_0) = \int_0^\infty c(y, \varphi(y)) d\tau$.

The refined fluid model (3.3) is resulted in by the following change of time scale through

$$\frac{dt}{d\tau} = \lambda(y, \varphi(y)) + \mu(y, \varphi(y)), \quad t(0) = 0,$$

where we recall that t is the time scale for the refined fluid model, and τ is the original time scale for the standard fluid model. If μ and λ are not separated from zero, then it might be the case that as τ goes to infinity, t does not. In this case, problem (3.3) differs from problem (3.2). This is illustrated in Example 3.1 below. As Example 3.1 shows, a policy optimal for (3.2) could be also not optimal for (3.3).

The above described change of time scale is similar to the uniformization technique (see Puterman (1994)) in the theory of CTMDPs, which transforms a continuous-time problem to its equivalent discrete-time version. Interestingly enough, while in the stochastic problem, this technique requires the transition rates to be bounded; in the fluid problem, to make sure that the transformation results in an equivalent problem, we see that what matters seems to be whether the transition rates are separated from zero, rather than their boundedness. More about the motivations for this change of time scale can be found after Conditions 3.1.

From now on in this chapter, by fluid approximation we always mean the refined fluid model (3.3). To our best knowledge, the quest for a refined fluid model receives no attention in the current literature.

Problem statement: Given the relations between the initial states in problems (3.1) and (3.3) to be "$Y_0 = ny_0$, suppose that $\varphi^*(y)$ is (nearly) optimal for problem (3.3). Then the aim of the current work is to prove that, the policy $\Phi^*(i) = \varphi^*(i/n)$ is nearly optimal for (3.1), if n is big; and to obtain explicit expressions for the accuracy of the aforementioned fluid approximations, in terms of the objective functional "$W^{\Phi}(i)$.

Conditions, which guarantee that problems (3.1) and (3.3) are well defined, are imposed in the next section.

3.3 Conditions and preliminary results

Conditions 3.1 *(a) For all $y > 0$ and $a \in A$,*

$$\lambda(y,a) \geq 0, \quad \mu(y,a) > 0, \quad \inf_{a\in A, y>0} \frac{\mu(y,a)}{\lambda(y,a)} \geq \tilde{\eta} \in (1,\infty).$$

(b) There exist constants $\eta_1 \in (1,\tilde{\eta})$, $\eta_2 \in (1,\frac{\tilde{\eta}}{\eta_1})$, $C_1 \in (0,\infty)$ and $C_2 \in (0,\infty)$ such that

$$\sup_{a\in A, y>0} \frac{|c(y,a)|}{[\lambda(y,a)+\mu(y,a)]\eta_1^y} \leq C_1 < \infty; \quad \sup_{a\in A, y>0} \frac{\lambda(y,a)+\mu(y,a)}{\eta_2^y} \leq C_2 < \infty.$$

Under Conditions 3.1, the change of time scale explained after formula (3.4) is motivated by the following observation: we note that

$$-1 \leq \frac{\lambda(y,a)-\mu(y,a)}{\lambda(y,a)+\mu(y,a)} \leq \frac{1-\tilde{\eta}}{1+\tilde{\eta}} < 0,$$

from which we conclude that the absorbing state $y = 0$ is reached in finite time (in scale t) from any initial state $y(0) > 0$, which is consistent with that the underlying stochastic model gets absorbed at state zero *a.s.*, see Lemma 3.1 below. In fact, under this observation, one can use any appropriate positive function f to define the change of time $\frac{dt}{d\tau} = f(y)$: we only want to be sure that $\lim_{t\to\infty} y(t) = 0$. See also the paragraph after Remark 3.1.

Definition 3.1 (Normal policy) *A feedback control policy $\varphi(y)$ in problem (3.3) is called normal, if the following function*

$$y \mapsto p(y) = \frac{c(y,\varphi(y))}{\mu(y,\varphi(y))-\lambda(y,\varphi(y))}, \tag{3.5}$$

defined on $\mathbb{R}_+^0 \setminus \{0\}$, is piece-wise Lipschitz continuous, i.e., there exist finite intervals

$(y_0 \triangleq 0, y_1), (y_1, y_2), \ldots$ *with* $\lim_{j\to\infty} y_j = \infty$, *such that on each of them,* $p(y)$ *is Lipschitz continuous.*

The existence of a normal policy is not always obvious. However, under Conditions 3.1 and Conditions 3.2, where the latter are imposed below, its existence is guaranteed, see Lemma 3.3.

Conditions 3.2 *There exist a constant* $\delta > 0$ *and finite intervals* $(y'_0 \triangleq 0, y'_1), (y'_1, y'_2), \ldots$ *with* $\lim_{j\to\infty} y'_j = \infty$, *such that on each of them,* $\inf_{y\in(y'_j, y'_{j+1}), a\in A}[\mu(y,a)+\lambda(y,a)] \geq \delta > 0$, *and functions* μ, λ *and* c *are Lipschitz continuous with respect to* y *for each fixed* $a \in A$, *and the Lipschitz constants are a-independent (but can be different for different intervals).*

Let us compare our conditions with closely related works.

Mandelbaum and Pats (1995) studied the trajectory-wise convergence for uncontrolled Birth-and-Death processes. Therefore, μ and λ are only state-dependent. By assuming that μ and λ are locally Lipschitz, and there exists a constant $K \geq 0$ such that $\forall y \geq 0$, $\lambda(y) \leq K(1+y)$, the authors showed that if ${}^nY_0 = ny(0)$, then as $n \to \infty$, $\frac{{}^nY_\tau}{n} \to y(\tau)$ *a.s.* and uniformly on $[0,T]$, where $T \geq 0$ can be arbitrarily fixed, see (Mandelbaum and Pats 1995, Sec.4.4, Thm.4.1). Here, $y(\tau)$ is from the standard fluid model (3.2). Compared with their conditions, Conditions 3.1 are weaker in that for any fixed $a \in A$, firstly, we do not require $\mu(y,a)$ and $\lambda(y,a)$ to be continuous in y; and secondly, we allow that $\lambda(y,a)$ and $\mu(y,a)$ can increase in y faster than linearly. On the other hand, since Mandelbaum and Pats (1995) was not only restricted to the absorbing case, the condition of $\frac{\mu(y)}{\lambda(y)} > 1$ was not required there. If one takes $\mu(y,a)$, $\lambda(y,a)$ and $c(y,a)$ to be a-independent, then our Conditions 3.1 are reduced to Conditions 3(b) in Piunovskiy (2009b), see also the discussion in Section 3.6 below. However, the results derived in (Piunovskiy 2009b, Sec.3) additionally require μ, λ and c, which are only y-dependent, to be piece-wise continuously differentiable.

Conditions similar to Conditions 3.2 were imposed in Pang and Day (2007), though the models studied there are different.

We emphasize that under Conditions 3.2, $\forall a \in A$, $\mu(y,a)$ and $\lambda(y,a)$ are not necessarily continuous in y: they only need be Lipschitz continuous on specific open intervals.

Lemma 3.1 *Suppose Conditions 3.1 are satisfied. Then problems (3.1) and (3.3) are well defined in the following sense:*
(a) For any Φ, *the controlled process* $\{{}^nY_t, t \geq 0\}$ *is regular.*
(b) For any Φ *and* ni, $|{}^nW^{\Phi}(ni)| < \infty$.
(c) For any φ *and* y, $|v^{\varphi}(y)| < \infty$, *where* $v^{\varphi}(y)$ *is given in (3.4).*

Let us remind that the Bellman equation for problem (3.3) is as follows:

$$\inf_{a\in A}\left\{\frac{\frac{dv}{dy}[\lambda(y,a)-\mu(y,a)]+c(y,a)}{\lambda(y,a)+\mu(y,a)}\right\}=0,\ v(0)=0. \tag{3.6}$$

Lemma 3.2 *Suppose Conditions 3.1 are satisfied. Then the following assertions hold:*
(a) Function

$$y\mapsto \inf_{a\in A}\left[\frac{c(y,a)}{\mu(y,a)-\lambda(y,a)}\right] \tag{3.7}$$

is bounded on each bounded interval $y\in(0,\hat{y}]$, and function

$$v(y)\triangleq\int_0^y \inf_{a\in A}\left[\frac{c(z,a)}{\mu(z,a)-\lambda(z,a)}\right]dz \tag{3.8}$$

satisfies the Bellman equation (3.6).
(b) If additionally, Conditions 3.2 are satisfied, then on each interval (y'_i,y'_{i+1}), $i=0,1,\ldots$, function

$$F(y,a)\triangleq\frac{c(y,a)}{\mu(y,a)-\lambda(y,a)} \tag{3.9}$$

is Lipschitz continuous with respect to y, $\forall a\in A$. Here, we remind that the concerned intervals come from Conditions 3.2. Moreover, the Lipschitz constants (on different intervals) are a-independent.

Remark 3.2 *Actually, function (3.7) is universally measurable, but integral (3.8) and other similar formulae are well defined, if we consider their Borel-measurable modifications, see (Bertsekas and Shreve 1978, Lem.7.27, Prop.7.47). In this connection, equality $\frac{dv}{dy}=\inf_{a\in A}\left[\frac{c(y,a)}{\mu(y,a)-\lambda(y,a)}\right]$ holds a.e. with respect to Lebesgue measure dy. In what follows, the values of $\frac{dv}{dy}$ on any set of (Lebesgue) measure 0 do not play any role.*

Lemma 3.3 *Under Conditions 3.1, suppose there exist finite intervals $(y'_0=0,y'_1)$, (y'_1,y'_2), $\ldots$ with $\lim_{i\to\infty}y'_i=\infty$, such that on each of them, $F(y,a)$ given in (3.9) is Lipschitz continuous (with respect to y), for each $a\in A$; and the Lipschitz constants are a-independent. Then $\forall\hat{y}>0$, $\forall\varepsilon>0$ there exists a normal policy φ^* in problem (3.3) on $(0,\hat{y}]$, such that*

$$\frac{dv^{\varphi^*}}{dy}=\frac{c(y,\varphi^*(y))}{\mu(y,\varphi^*(y))-\lambda(y,\varphi^*(y))}\le\inf_{a\in A}\frac{c(y,a)}{\mu(y,a)-\lambda(y,a)}+\varepsilon=\frac{dv}{dy}+\varepsilon, \tag{3.10}$$

where the first equality holds a.e. and the first inequality holds $\forall\, y>0$; and $\forall y\in(0,\hat{y}]$,

$$v(y)\le v^{\varphi^*}(y)\le v(y)+\varepsilon\hat{y}, \tag{3.11}$$

where v is given in (3.8).

Evidently, the conditions in Lemma 3.3 are satisfied if Conditions 3.1 and 3.2 hold, see part (b) of Lemma 3.2.

3.4 Main results

Theorem 3.1 *Suppose Conditions 3.1 are satisfied, φ^* is a normal feedback control policy in problem (3.3) on $(0,\hat{y}+1]$, points $y_1,y_2,\dots,y_L<\hat{y}+1\le y_{L+1},\dots$ come from Definition 3.1, and D is the common Lipschitz constant of function (3.5) with $\varphi=\varphi^*$ on all of the intervals (y_j,y_{j+1}), $j=0,1,\dots,L$. Then, for*

$$\Phi^*(i)=\varphi^*(i/n):\ S\to A, \tag{3.12}$$

the following assertions hold:

(a)

$$\sup_{0\le i\le\hat{y}n}\left|{}^{n}W^{\Phi^*}(i)-v^{\varphi^*}(i/n)\right|\le\frac{\hat{\varepsilon}(n)}{2}. \tag{3.13}$$

Here and below, we put

$$\hat{\varepsilon}(n)\triangleq\frac{2K_1}{n}+\frac{2K_2}{\tilde{\eta}^n}+2K_3(\eta_1^{1/n}-1),$$

$$K_1\triangleq\frac{\tilde{\eta}+1}{\tilde{\eta}-1}[D(\hat{y}+1)+3C_1L\eta_1^{\hat{y}+1}],$$

$$K_2\triangleq\frac{\tilde{\eta}+1}{\tilde{\eta}-1}C_1\left[1+\frac{2(\tilde{\eta}+1)}{(\tilde{\eta}-1)\ln\eta_1}\right]\frac{\eta_1^{\hat{y}+1}\tilde{\eta}^2}{\tilde{\eta}-\eta_1},$$

$$K_3\triangleq\left(\frac{\tilde{\eta}+1}{\tilde{\eta}-1}\right)^2\frac{3C_1L\eta_1^{\hat{y}+1}}{\ln\eta_1}.$$

(b)

$$\sup_{0\le i\le\hat{y}n}\left|{}^{n}W^{\Phi^*}(i)-\inf_{\Phi}{}^{n}W^{\Phi}(i)\right|\le\sup_{y\in(0,\hat{y}+1]}\left[\frac{dv^{\varphi^*}}{dy}-\frac{dv}{dy}\right]\times\frac{(\hat{y}+1)(\tilde{\eta}+1)}{\tilde{\eta}-1}+\hat{\varepsilon}(n). \tag{3.14}$$

Remark 3.3 *Function $\left|\frac{dv^{\varphi^*}}{dy}\right|=\frac{|c(y,\varphi^*(y))|}{\mu(y,\varphi^*(y))-\lambda(y,\varphi^*(y))}$ is bounded by $C_1\eta_1^y\frac{\tilde{\eta}+1}{\tilde{\eta}-1}$ according to Conditions 3.1; similarly, function $\left|\frac{dv}{dy}\right|$ is also bounded on bounded intervals, see also Remark 3.2. Remember also that $\frac{dv^{\varphi^*}}{dy}\ge\frac{dv}{dy}$ for any normal feedback policy φ^*, see (3.4) and (3.8).*

Corollary 3.1 *Suppose all the conditions of Theorem 3.1 are satisfied, and inequality (3.10) holds for all $y\in(0,\hat{y}+1]$. Then, for large enough n, $\Phi^*(i)=\varphi^*(i/n)$ is nearly* *optimal for all initial states* $^nY_0 \in [0, \hat{y}n]$ *in problem (3.1). Namely,*

$$\sup_{0 \le i \le \hat{y}n} \left| {}^nW^{\Phi^*}(i) - \inf_{\Phi} {}^nW^{\Phi}(i) \right| \le \delta + \hat{\varepsilon}(n), \ \forall n = 1, 2, \ldots,$$

where $\delta = \frac{\varepsilon(\hat{y}+1)(\hat{\eta}+1)}{\hat{\eta}-1}$, *with* ε *coming from (3.10). We remind that* $\lim_{n\to\infty} \hat{\varepsilon}(n) = 0$, *see Theorem 3.1.*

Obviously, if Conditions 3.1 and Conditions 3.2 are satisfied, then the statements of Theorem 3.1 and Corollary 3.1 hold, and $\forall \hat{y}$, the value of $\varepsilon > 0$ (and hence $\delta > 0$) appearing in Corollary (3.1) can be chosen arbitrarily small, see Lemma 3.3. At the same time, constants K_1 and K_3 depending on φ^* can depend on ε, simply because φ^* can depend on ε, see Lemma 3.3. In any case, after the values of ε, K_1 and K_3 are fixed, $\lim_{n\to\infty} \hat{\varepsilon}(n) = 0$.

3.5 Examples

Example 3.1 *Let A be a singleton, that is, we consider an uncontrolled case. Therefore, the argument a is omitted everywhere. Suppose*

$$\begin{aligned}
\lambda(y) &= I\{y \in (0,1]\} + \frac{1}{2}(y-1)^2 I\{y \in (1,3]\} + 2I\{y \in (3,\infty)\}, \\
\mu(y) &= 2I\{y \in (0,1]\} + (y-1)^2 I\{y \in (1,3]\} + 4I\{y \in (3,\infty)\}, \\
c(y) &= 3I\{y \in (0,1]\} + \frac{3}{2}(y-1)^2 I\{y \in (1,3]\} + 6I\{y \in (3,\infty)\},
\end{aligned}$$

so that Conditions 3.1 are satisfied, function $\frac{c(y)}{\mu(y)-\lambda(y)} \equiv 3$ *is (globally) Lipschitz, and* $v(y) = 3y$. *Suppose* $\frac{^nY_0}{n} = y(0) = 2$.

Solution: If the standard fluid model (3.2) is accepted, then we see that

$$y(\tau) = \frac{2}{\tau+2} + 1 \to 1 \text{ as } \tau \to \infty.$$

Hence

$$\int_0^\infty c(y(\tau))d\tau = 3 \ne v(2),$$

because $\{y(\tau), \tau \ge 0\}$ is never absorbed at state zero.

It is interesting to compare $\{y(\tau), \tau \ge 0\}$ with the corresponding stochastic process $\{^nY_u, u \ge 0\}$, starting from $^nY_0 = 2n$. According to (Mandelbaum and Pats 1995, Thm.4.1) (see also the paragraph after Conditions 3.2, trajectories of $^nY_\tau/n$ converge *a.s.* to $y(\tau)$ uniformly on each fixed finite interval $[0, T]$, as $n \to \infty$. Consequently, for the performance functional similar to (3.1) we have

$$\lim_{n\to\infty} E_{2n}\left[\int_0^T {}^nC({}^nY_u)du\right] = \int_0^T c(y(\tau))d\tau = 3 - \frac{6}{T+2},$$

whose proof is based on Lebesgue dominated convergence theorem. Therefore,

$$\lim_{T\to\infty}\lim_{n\to\infty} E_{2n}\left[\int_0^T {}^nC({}^nY_u)du\right] = 3. \tag{3.15}$$

But we are interested in the expected total cost at large values of n, that is, in the following limit:

$$\lim_{n\to\infty}\lim_{T\to\infty} E_{2n}\left[\int_0^T {}^nC({}^nY_u)du\right] = \lim_{n\to\infty} {}^nW(2n), \tag{3.16}$$

and this quantity is far different from 3. Indeed, according to Theorem 3.2,

$$\lim_{n\to\infty} {}^nW(2n) = v(2) = 6.$$

Note that as n increases, the absorption of $\{{}^nY_u/n, u \geq 0\}$ at zero is postponed to ∞ (time) in the sense that for any arbitrarily fixed $T \geq 0$, $\frac{{}^nY_\tau}{n} \to y(\tau)$ *a.s.* uniformly on $[0,T]$ as $n \to \infty$, and $y(\tau) \to 1$ as $\tau \to \infty$. This observation is confirmed by computer simulations, see Figure 3.1. In this connection, one may view formula (3.15) as it corresponds to the 'absorption' of the process $\{{}^nY_u/n, u \geq 0\}$ at state one. On the opposite, formula (3.16) provides the expected total cost up to the absorption at zero: recall that for any fixed n, $\lim_{u\to\infty} {}^nY_u = 0$ *a.s.* (see Lemma 3.1).

On the other hand, if we accept the refined fluid model (3.3), then $\frac{dy}{dt} = -1/3$, and the absorption of the refined fluid model at zero occurs in finite time. Therefore, if $y(0) = 2$, $\int_0^\infty \frac{c(y(t))}{\lambda(y(t))+\mu(y(t))}dt = \int_0^6 dt = v(2) = 6$.

By the way, $\frac{dt}{d\tau} = \lambda(y(\tau)) + \mu(y(\tau)) = \frac{6}{(\tau+2)^2}$ and $t(\tau) = 3 - \frac{6}{\tau+2}$, so that $\lim_{\tau\to\infty} t(\tau) = 3$. ■

In this example, we see that if $\{y(\tau), \tau \geq 0\}$ does not get absorbed at zero, then it can happen that the standard fluid model provides inaccurate approximations to the scaled stochastic model. The system considered there is an uncontrolled one. Now let us consider the controlled problem (3.2), and suppose that φ^* is optimal for it, under which $\lim_{\tau\to\infty} y(\tau) = \kappa > 0$, and $y(0) > \kappa$. Then Φ^* given by $\Phi^*(i) = \varphi^*(i/n)$ can easily be not nearly optimal for the scaled stochastic problem (3.1), simply because the values of $\varphi^*(y)$ at $y < \kappa$ play no role for the standard fluid model, and can thus be taken arbitrarily.

Remark 3.4 *The detailed calculations involved in Example 3.2 and Example 3.3 below are quite tedious. Consequently, we include them at the end of Section 3.7.*

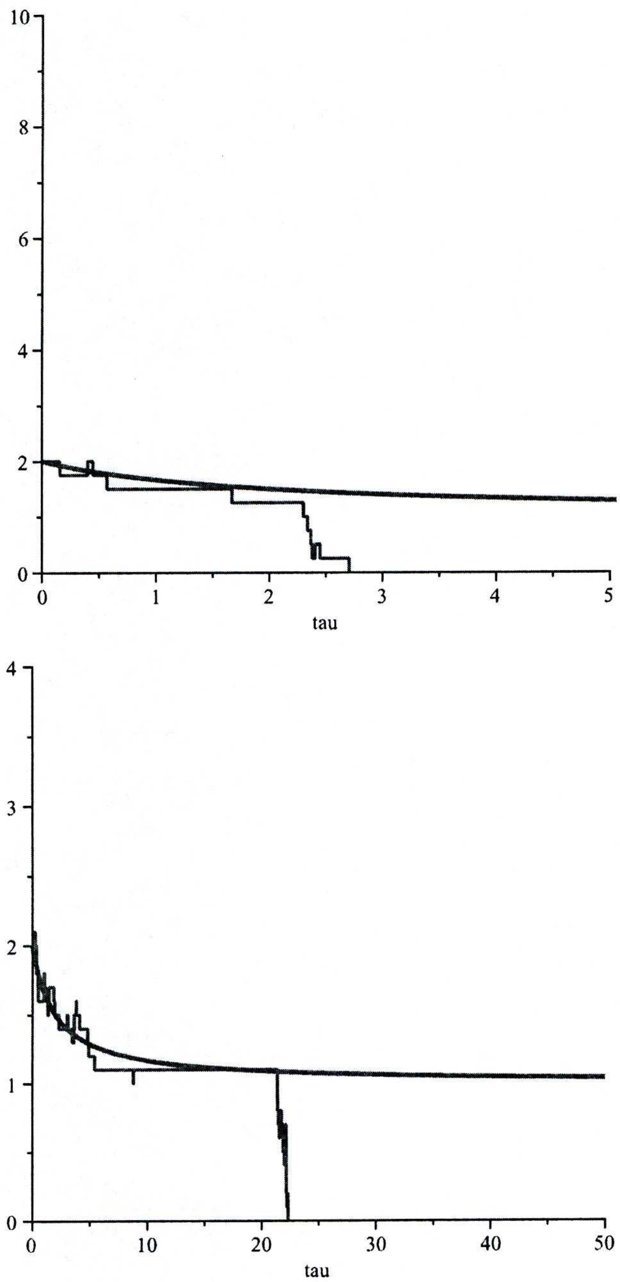

Figure 3.1: Trajectories of the fluid model $y(\tau)$ and of the stochastic system $\frac{{}^{n}Y_{\tau}}{n}$ for $n = 4$ and $n = 10$.

Example 3.2 *Consider an M/M/1 queueing system with a controlled input stream*

$$n\lambda(j/n,a) = n(d_0 + d_1 a),$$

$d_0, d_1 > 0$, $d_0 + d_1 < 1$, $a \in A = [0,1]$. *The service intensity*

$$n\mu(j/n,a) = n$$

is constant. As usual, n is a fixed scaling parameter. The initial state is $i > 0$. Suppose we are interested in the expected total throughput as well as the expected total queue length. In general, one cannot maximize the expected throughput and minimize the expected total queue length at the same time: there is always a trade-off to be made. Here we regard both of the criteria important, and therefore, we take the following utility function

$$^{n}C(i,a) = c(i/n,a) = i/n - Ra,$$

where $R > 0$ is a given constant representing the relative importance of the throughput. One has to solve problem (3.1). The corresponding fluid model is defined by

$$\lambda(y,a) = d_0 + d_1 a, \; \mu(y,a) = 1, \; c(y,a) = y - Ra.$$

Here and above in this example, $\lambda(0,a) = \mu(0,a) = c(0,a) \equiv 0$. Note that Conditions 3.1 are satisfied.

Solution: To study the fluid model, we solve the Bellman equation (3.6) and obtain

$$v(y) = \begin{cases} \dfrac{Ry - y^2/2}{d_0 + d_1 - 1}, & \text{if } 0 \le y \le y^*; \\[2ex] \dfrac{y^2 - (y^*)^2}{2(1-d_0)} - \dfrac{Ry^* - (y^*)^2/2}{1 - d_0 - d_1}, & \text{if } y > y^*, \end{cases}$$

where

$$y^* \triangleq \frac{R(1-d_0)}{d_1}.$$

Feedback control policy

$$\varphi^*(y) \triangleq \begin{cases} 1, & \text{if } 0 < y \le y^*; \\ 0, & \text{if } y > y^* \end{cases}$$

is normal and optimal for problems (3.2) and (3.3), and $\varepsilon = 0$ in formula (3.10).

For any fixed n, one can also explicitly solve the stochastic optimal control problem (3.1). The Bellman function $^{n}V(i) = \inf_{\Phi} {}^{n}W^{\Phi}(i)$, satisfying the following Bellman equation

$$\inf_{a\in A}\left\{{}^nC(i,a)+n\lambda(\frac{i}{n},a)\,{}^nV(i+1)+n\mu(\frac{i}{n},a)\,{}^nV(i-1)\right.$$
$$\left.-n[\lambda(\frac{i}{n},a)+\mu(\frac{i}{n},a)]\,{}^nV(i)\right\}=0,\ \text{if } i>0;$$
$$V(0)=0,$$

is given by the following formulae: if $0\le i\le i^*\stackrel{\triangle}{=}\left[\frac{nR(1-d_0)}{d_1}-\frac{1+d_0}{2(1-d_0)}-\frac{1}{2}\right]$, then

$$\begin{aligned}{}^nV(i) &= \left[\frac{2i^*+1-2nR}{1-d_0-d_1}+\frac{d_0+d_1+1}{(1-d_0-d_1)^2}-\frac{2i^*+1}{1-d_0}-\frac{1+d_0}{(1-d_0)^2}\right]\\ &\times\frac{(d_0+d_1)^{i^*+1}-(d_0+d_1)^{i^*-i+1}}{2n^2(1-d_0-d_1)}\\ &+\frac{1}{2n^2}\left[\frac{i^2}{1-d_0-d_1}+i\left(\frac{d_0+d_1+1}{(1-d_0-d_1)^2}-\frac{2nR}{1-d_0-d_1}\right)\right];\end{aligned}$$

if $i>i^*$, then

$$\begin{aligned}{}^nV(i) &= \left[\frac{2i^*+1-2nR}{1-d_0-d_1}+\frac{d_0+d_1+1}{(1-d_0-d_1)^2}-\frac{2i^*+1}{1-d_0}-\frac{1+d_0}{(1-d_0)^2}\right]\\ &\times\frac{(d_0+d_1)^{i^*+1}-d_0-d_1}{2n^2(1-d_0-d_1)}+\frac{1}{2n^2}\left[(i^*)^2\left(\frac{1}{1-d_0-d_1}-\frac{1}{1-d_0}\right)\right.\\ &+i^*\left(\frac{d_0+d_1+1}{(1-d_0-d_1)^2}-\frac{2nR}{1-d_0-d_1}-\frac{1+d_0}{(1-d_0)^2}\right)\\ &\left.+\frac{i^2}{1-d_0}+\frac{i(1+d_0)}{(1-d_0)^2}\right].\end{aligned}$$

Policy

$$\hat{\Phi}(i)\stackrel{\triangle}{=}\begin{cases}1, & \text{if } 0<i\le i^*;\\ 0, & \text{if } i>i^*\end{cases}$$

is optimal in problem (3.1).

Clearly, $i^*/n\to y^*$ as $n\to\infty$, so that policies in the stochastic model $\hat{\Phi}(i)$ and $\Phi^*(i)=\varphi^*(i/n)$ are close to each other. Theorem 3.1 says that

$$v(i/n)=v^{\varphi^*}(i/n)\approx {}^nW^{\Phi^*}(i)$$

and

$${}^nW^{\Phi^*}(i)\approx {}^nW^{\hat{\Phi}}(i)={}^nV(i).$$

Let us fix $d_0=0.25$, $d_1=0.5$ and $R=1$. The graphs of $v(i/n)$ and ${}^nV(i)$ for $n=10$ and $n=100$ are presented in Figure 3.2. For these values of parameters, one can take $\tilde{\eta}=4/3$, $\eta_1=7/6$, $\eta_2=15/14$, $C_1=2.23$, $C_2=1.75$. For the normal policy φ^*,

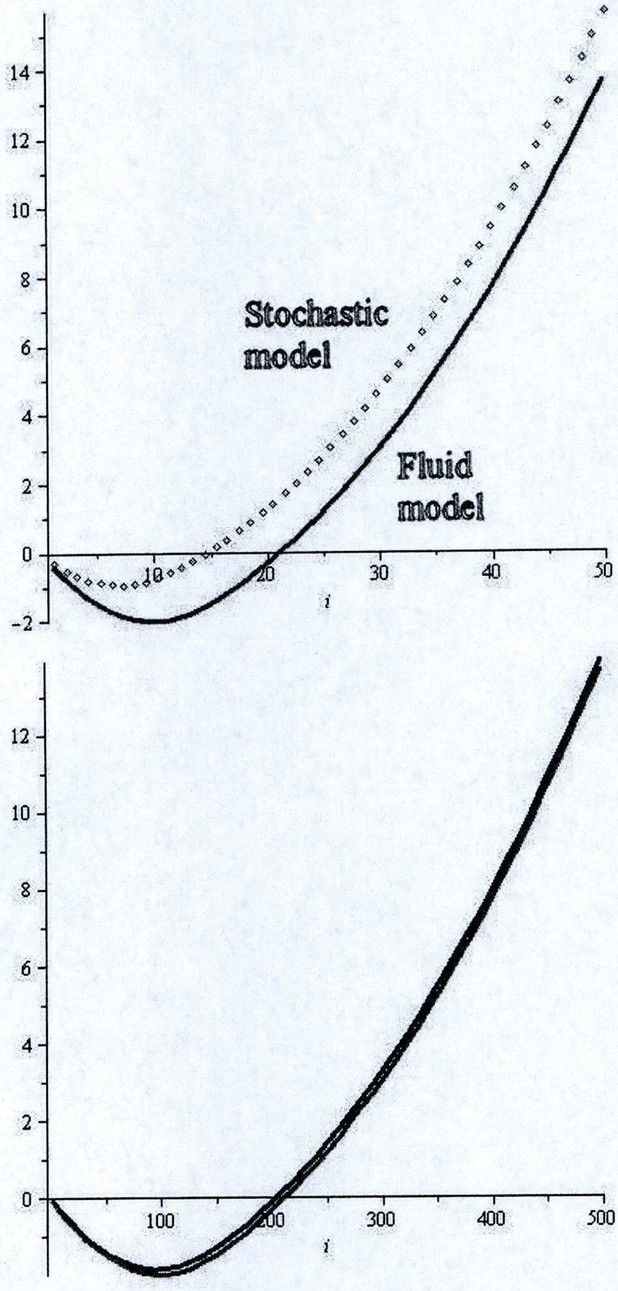

Figure 3.2: Graphs of Bellman functions (Example 3.2): $n = 10$ on the top and $n = 100$ on the bottom.

$y^* = 1.5$, $L = 1$, $D = 4$. Now, after choosing $\hat{y} = 5$, we have $K_1 = 286$, $K_2 = 38500$, $K_3 = 5360$, and for $n = 100$ we obtain 11. and 22. on the righthand side of (3.13) and (3.14) correspondingly, meaning that our theoretical estimate is very rough and the real accuracy is 100 times more precise. For $n \geq 100{,}000$ the presented numbers change to 0.011 and 0.022; the accuracy is (approximately) inversely proportional to the value of n. ■

Example 3.3 *Consider the optimal control problem in the same set-up as in Example 2, except for the service intensity, now modelling the case when there are infinitely many servers. If we put $\mu(y,a) = y$, then state zero in problem (3.2) is no longer "absorbing": for small enough $y > 0$, $\lambda(y,a) - \mu(y,a) > 0$. To avoid this and make the example consistent with the previous material, we say that the 'transient' stage (nearly) finishes as soon as $y = 1$. Then after the change of variable $y - 1 :\rightarrow y$, we introduce the following parameters:*

$$\lambda(y,a) = d_0 + d_1 a, \quad \mu(y,a) = y + 1, \quad c(y,a) = y + 1 - Ra$$

with the absorbing state zero. As in Example 3.2, $\lambda(0,a) = \mu(0,a) = c(0,a) \equiv 0$. The formulae for the parameters of the stochastic model are obvious. One can check that Conditions 3.1 are satisfied.

Solution: The most interesting case is when $R \in (d_1, d_1/(1-d_0))$; otherwise either $\varphi^*(y) \equiv 0$ or $\varphi^*(y) \equiv 1$ solves problems (3.2) and (3.3). We put $y^* \stackrel{\triangle}{=} \frac{d_1 - R(1-d_0)}{R - d_1}$. Now the Bellman equation (3.6) has the following solution

$$v(y) = \begin{cases} y + d_0 \ln\left(\dfrac{y+1-d_0}{1-d_0}\right), & \text{if } 0 \leq y \leq y^*; \\[2ex] y + (d_0 + d_1 - R)\ln\left(\dfrac{y+1-d_0-d_1}{y^*+1-d_0-d_1}\right) & \\[2ex] +d_0 \ln\left(\dfrac{y^*+1-d_0}{1-d_0}\right) & \text{if } y > y^*. \end{cases}$$

Policy

$$\varphi^*(y) \stackrel{\triangle}{=} \begin{cases} 0, & \text{if } 0 < y \leq y^*; \\ 1, & \text{if } y > y^* \end{cases}$$

is normal and optimal for problems (3.2) and (3.3), and $\varepsilon = 0$ in formula (3.10).

Let us fix $d_0 = 0.25$, $d_1 = 0.5$, $R = 7/12$. For these values of parameters, $\tilde{\eta} = 4/3$, $\eta_1 = 7/6$, $\eta_2 = 15/14$, $C_1 = 4$, $C_2 = 6.121$. For the normal policy φ^*, $y^* = 0.75$, $L = 1$, $D = 4$. Now, after choosing $\hat{y} = 5$, we have $K_1 = 380$, $K_2 = 69200$, and $K_3 = 9620$. If $n \geq 100{,}000$, then on the righthand side of (3.13) and (3.14) we obtain 0.019 and 0.038, correspondingly. Note that function $v(y)$ ranges over $[0,5]$ if $y \in [0, \hat{y} = 5]$, meaning that the fluid approximation is fairly accurate. ■

3.6 Comments and conclusion

Comments on the obtained estimate for the rate of convergence: Example 3.2 and Example 3.3 show that the estimate for the rate of convergence, $\hat{\varepsilon}(n) = constant \times O(1/n)$ is tight, as far as the order is concerned. The order of $\frac{1}{n}$ for the convergence is also observed in Example 4 in Piunovskiy (2009b). In this sense, we claim that our estimate is accurate. On the other hand, as revealed by Example 3.2, the estimated accuracy of fluid approximations can still be much rougher than the actual one, meaning that our estimate for the constant can be very crude (the constant we mean comes from $\hat{\varepsilon}(n) = constant \times O(1/n)$). However, if one aims at working on general conditions covering a broad class of models, then it seems difficult to obtain an estimate accurate for both the order and the constant.

In fact, (Piunovskiy 2009b, Ex.4) is the same as our Example 3.2, except for the cost rate c, which is bounded there and unbounded here. The parameters in (Piunovskiy 2009b, Ex.4) also satisfy our Conditions 3.1, while the parameters in our Example 3.2 do not satisfy Conditions 3(b) assumed there.

If one fixes a (feedback) control policy φ^*, then in the case of $\lambda(y,\varphi^*(y)) > 0$, Conditions 3.1 are essentially reduced to Conditions 3(b) in Piunovskiy (2009b). Then according to (Piunovskiy 2009b, Lem.3, Lem.4), one can construct an exponential Lyapunov function (see (Piunovskiy 2009b, Con.1)), and if additionally, $\lambda(y,\varphi^*(y))$, $\mu(y,\varphi^*(y))$ and $c(y,\varphi^*(y))$ are piece-wise continuously differentiable, we have (Piunovskiy 2009b, Thm.1) validated, which together with (Piunovskiy 2009b, Lem.4), also implies the convergence result like part (a) of Theorem 3.1. However, it did not provide any estimate for the rate of convergence. Although under the extra conditions requiring $\lambda(y,\varphi^*(y))$, $\mu(y,\varphi^*(y))$ and $c(y,\varphi^*(y))$ to be continuously differentiable with uniformly bounded first order derivatives, an estimate for the rate of convergence was provided by (Piunovskiy 2009b, Cor.1), which is also of order $\frac{1}{n}$; such extra conditions seem quite strong and restrictive: under φ^* derived in Example 3.2, the parameters there do not satisfy them.

Comments on the methods of proof: In Section 3.1, we mentioned two possible approaches to justify fluid approximations to Markovian optimal control problems: the first approach is based on the convergence of trajectories, and the second one is based on the study of algebraic equations of the dynamic programming type. This chapter follows the second approach and is based on the Dynkin's formula, like in the proof of Theorem 2 in Piunovskiy (2009b), which allows us to obtain an estimate for the rate of convergence (performance functional-wise). On the other hand, the proof of Theorem 1 in Piunovskiy (2009b) does not invoke the Dynkin's formula, but is based on the numerical methods for differential equations: in the uncontrolled case (or if a control policy is fixed) equation

$$c\left(\frac{i}{n}\right)+n\lambda\left(\frac{i}{n}\right){}^{n}W(i+1)+n\mu\left(\frac{i}{n}\right){}^{n}W(i-1)-n\left(\lambda\left(\frac{i}{n}\right)+\mu\left(\frac{i}{n}\right)\right){}^{n}W(i)=0$$

(see (3.1)) is simply a version of the Euler scheme for the differential equation

$$c(y)+\frac{dv}{dy}(\lambda(y)-\mu(y))=0.$$

The main reason for following the second approach in the current work lies in that the underlying problem is on an infinite horizon: from Example 3.1, we already see that the first approach can be inefficient, simply because bad things can happen as time goes to infinity. The first approach seems more applicable, when the underlying horizon is finite and if one aims at estimating the rate of convergence. For example, by following the first approach, for a bandwidth-sharing network, Chapter 5 below will establish the rate of convergence (performance functional-wise), which is of order $\frac{1}{\sqrt{n}}$. However, there, it seems that the rate of convergence (in terms of the order) can vary for different values of parameters in the model. Indeed, by following the second approach and using the Dynkin's formula, with some extra conditions on the parameters, we shall also obtain an estimate of order $\frac{1}{n}$.

Other general comments: Absorbing (at zero) Birth-and-Death processes are closely related to EOQ (Economic Order Quantity) models in Inventory theory, and results like those derived in the current work and Piunovskiy (2009b) are rather useful for the studies of fluid approximations there, where the performance functional, the so called TCU, is a long-run average, and this will be demonstrated in Chapter 4. This chapter is about continuous-time Birth-and-Death processes only. Nevertheless, there is no doubt that the same results are valid also for discrete-time random walks: the one-to-one correspondence between those models is well known and widely used in Piunovskiy (2009b).

Conclusions: While investigating stochastic optimal control problems, one is concerned about the performance functionals. Therefore, in such situations, if one is interested in the accuracy of fluid approximations, then the straightforward study of equations of the dynamic programming type seems more reasonable than the analysis of particular trajectories. Quite formally, this approach is close to approximations of differential operators and to numerical schemes for ordinary and partial differential equations. These matters were discussed in Piunovskiy (2009b), see the proof of (Piunovskiy 2009b, Thm.1). In this chapter, like in the proof of (Piunovskiy 2009b, Thm.2), the proofs are based on the Dynkin's formula and lead to explicit expressions for the accuracy of fluid approximations, in terms of the performance functional.

The suggested methods can be also applied to many other specific cases like fluid approximations to random walks, models with local transitions on multiple-dimensional lattices (Piunovskiy (2009a)), optimal control problems with long run average costs and with discounting, and so on. Possible applications include queueing systems, telecommunications, epidemics and inventories.

3.7 Proof of main statements

Proof of Lemma 3.1. (a) Under Conditions 3.1, for any $\eta > 1$, any value of n and all $i \geq 0$, the following relation holds:

$$n\lambda(i/n,a)\eta^{\frac{i+1}{n}} + n\mu(i/n,a)\eta^{\frac{i-1}{n}} - n[\lambda(i/n,a)+\mu(i/n,a)]\eta^{\frac{i}{n}} \leq 0 < \eta^{\frac{i}{n}},$$

see (Piunovskiy 2009b, (11)). Therefore, (Guo et al 2006, Assump.A) holds, following from which the controlled process $\{{}^nY_t, t \geq 0\}$ is regular under any control policy Φ.

(b) It follows immediately from the reasoning presented after Conditions 3.2 and (Piunovskiy 2009b, Lem.1, Lem.4).

(c) It follows immediately from Conditions 3.1 and expression (3.4). ■

Proof of Lemma 3.2. (a) The first part of this sub-statement is obvious. Indeed, if $y \in (0, \hat{y}]$, then

$$\begin{aligned}\left|\inf_{a\in A}\frac{c(y,a)}{\mu(y,a)-\lambda(y,a)}\right| &\leq \sup_{a\in A}\frac{|c(y,a)|}{\mu(y,a)+\lambda(y,a)} \times \sup_{a\in A}\frac{\mu(y,a)+\lambda(y,a)}{\mu(y,a)-\lambda(y,a)} \\ &\leq C_1\eta_1^{\hat{y}}\frac{\tilde{\eta}+1}{\tilde{\eta}-1}.\end{aligned}$$

As for the second part of this sub-statement, we shall indeed firstly prove a preliminary result, which we call the "incidental statement", asserting that under Conditions 3.1, $v(y)$ defined in (3.8) solves the following Bellman equation

$$\inf_{a\in A}\left\{\frac{dv}{dy}[\lambda(y,a)-\mu(y,a)]+c(y,a)\right\} = 0, \; v(0) = 0. \tag{3.17}$$

It is easy to see that $\forall a \in A$,

$$\begin{aligned}&\frac{dv}{dy}[\lambda(y,a)-\mu(y,a)]+c(y,a) \\ = \; &[\mu(y,a)-\lambda(y,a)]\left\{-\inf_{a\in A}\left\{\frac{c(y,a)}{\mu(y,a)-\lambda(y,a)}\right\}+\frac{c(y,a)}{\mu(y,a)-\lambda(y,a)}\right\} \geq 0,\end{aligned}$$

from which we conclude that

$$\inf_{a\in A}\left\{\frac{dv}{dy}[\lambda(y,a)-\mu(y,a)]+c(y,a)\right\} = 0,$$

since $0 < \mu(y,a)-\lambda(y,a) \leq C_2\eta_2^y$. Thus, the incidental statement is proved.

Now if one puts $\hat{\lambda} \stackrel{\triangle}{=} \frac{\lambda}{\lambda+\mu}$, $\hat{\mu} \stackrel{\triangle}{=} \frac{\mu}{\lambda+\mu}$ and $\hat{c} \stackrel{\triangle}{=} \frac{c}{\lambda+\mu}$, then equation (3.6) is exactly in the same form as (3.17). In addition, Conditions 3.1 are obviously satisfied by $\hat{\lambda}$, $\hat{\mu}$ and $\hat{c}$. Therefore, it follows from the incidental statement that $v(y)$ given in (3.8) also solves the Bellman equation (3.6), as required.

(b) Consider $F(y,a)$ on any fixed interval (y_i', y_{i+1}'), where we take $k > 0$ as the common Lipschitz constant (with respect to y) for μ, λ and c. Such a constant k can be taken because of Conditions 3.2. Since

$$1 \leq \frac{\mu(y,a)+\lambda(y,a)}{\mu(y,a)-\lambda(y,a)} \leq \frac{\tilde{\eta}+1}{\tilde{\eta}-1},$$

$$\sup_{a\in A, y\in(y_i', y_{i+1}')} [\lambda(y,a)+\mu(y,a)] \leq C_2\eta_2^{y_{i+1}'}$$

and

$$\sup_{a\in A, y\in(y_i', y_{i+1}')} |c(y,a)| \leq C_1\eta_1^{y_{i+1}'}C_2\eta_2^{y_{i+1}'},$$

we have for $y_1, y_2 \in (y_i', y_{i+1}')$,

$$\begin{aligned} |F(y_1,a)-F(y_2,a)| \leq & \left[\left(\frac{\tilde{\eta}+1}{\tilde{\eta}-1}\right)\Big/\delta\right]^2 [|c(y_1,a)(\mu(y_2,a)-\mu(y_1,a))| \\ & +|\mu(y_1,a)(c(y_1,a)-c(y_2,a))| \\ & +|c(y_1,a)(\lambda(y_2,a)-\lambda(y_1,a))| \\ & +|\lambda(y_1,a)(c(y_1,a)-c(y_2,a))|] \\ \leq & K|y_1-y_2|, \end{aligned} \tag{3.18}$$

where

$$K = 2C_2\eta_2^{y_{i+1}'}\left[C_1\eta_1^{y_{i+1}'}+1\right]\left[\left(\frac{\tilde{\eta}+1}{\tilde{\eta}-1}\right)\Big/\delta\right]^2 k > 0$$

can be taken as the a-independent Lipschitz constant under consideration. ■

Proof of Lemma 3.3. Consider a particular interval (y_i', y_{i+1}'). Then we know that there exists a Lipschitz constant K such that $\forall y_1, y_2 \in (y_i', y_{i+1}')$,

$$\left|\frac{c(y_1,a)}{\mu(y_1,a)-\lambda(y_1,a)} - \frac{c(y_2,a)}{\mu(y_2,a)-\lambda(y_2,a)}\right| \leq K|y_1-y_2|.$$

Let us put $\Delta \stackrel{\triangle}{=} \min\{\frac{\varepsilon}{3K}, \frac{y_{i+1}'-y_i'}{2}\}$, and fix an arbitrary $a_i^! \in A$ providing

$$\frac{c(y_i'+\Delta, a_i^!)}{\mu(y_i'+\Delta, a_i^!)-\lambda(y_i'+\Delta, a_i^!)} \leq \inf_{a\in A} \frac{c(y_i'+\Delta, a)}{\mu(y_i'+\Delta, a)-\lambda(y_i'+\Delta, a)} + \frac{\varepsilon}{3}.$$

Now if $y \in (y'_i, y'_{i+1})$ and $|y - (y'_i + \Delta)| \leq \Delta$, then $\forall a \in A$,

$$\frac{c(y,a_i^!)}{\mu(y,a_i^!) - \lambda(y,a_i^!)} \leq \frac{c(y'_i+\Delta,a_i^!)}{\mu(y'_i+\Delta,a_i^!) - \lambda(y'_i+\Delta,a_i^!)} + K\Delta$$

$$\leq \frac{c(y'_i+\Delta,a)}{\mu(y'_i+\Delta,a) - \lambda(y'_i+\Delta,a)} + \frac{\varepsilon}{3} + K\Delta$$
$$\leq \frac{c(y,a)}{\mu(y,a) - \lambda(y,a)} + K\Delta + \frac{2\varepsilon}{3},$$

and hence
$$\frac{c(y,a_i^!)}{\mu(y,a_i^!) - \lambda(y,a_i^!)} \leq \inf_{a\in A} \frac{c(y,a)}{\mu(y,a) - \lambda(y,a)} + \varepsilon.$$

Let us put $\varphi^*(y) = a_i^!$ for $y \in [y'_i, y'_i + 2\Delta)$. Now one can repeat the reasoning with $y'_i + 2\Delta$ instead of y'_i, given that $y'_i + 2\Delta < y'_{i+1}$. Then after a finite number of steps, we can eventually construct the required piece-wise constant mapping φ^* on $[y'_i, y'_{i+1})$. This reasoning can be applied to the other intervals, while if needed, one should modify the values of $\varphi^*(y'_i)$, $i = 1, 2, \ldots$. Finally, we finish with φ^* satisfying (3.10).

Evidently, function (3.5) is Lipschitz continuous on each of the intervals (y_j, y_{j+1}), where $\varphi^*(y)$ is constant. Obviously, $\lim_{j\to\infty} y_j = \infty$.

For the control policy φ^*, we remind that $v^{\varphi^*}(y) = \int_0^y \frac{c(z,\varphi^*(z))}{\mu(z,\varphi^*(z)) - \lambda(z,\varphi^*(z))} dz$, see (3.4). With in mind (3.8) for v, now it follows that $v^{\varphi^*}(y) \leq \int_0^y \inf_{a\in A} \frac{c(z,a)}{\mu(z,a)-\lambda(z,a)} dz + \varepsilon y \leq v(y) + \varepsilon \hat{y}$, as required. ■

Before we proceed to the proof of Theorem 3.1, let us remind several statements of technical nature (in the framework of Theorem 3.1). Most of them were proved in Piunovskiy (2009b), and the others can be easily proved in a similar way.

A1. Straightforward estimations of function (3.4) show that, for any control policy φ,
$$|v^{\varphi}(y)| \leq \sup_{z>0, a\in A} \left\{ \frac{\mu(z,a) + \lambda(z,a)}{\mu(z,a) - \lambda(z,a)} \right\} \int_0^y C_1 \eta_1^z dz \leq C_1 \frac{\eta_1^y - 1}{\ln \eta_1} \times \frac{\tilde{\eta}+1}{\tilde{\eta}-1}$$

and
$$|v^{\varphi}(y+d) - v^{\varphi}(y)| \leq C_1 \frac{\eta_1^y |\eta_1^d - 1|}{\ln \eta_1} \times \frac{\tilde{\eta}+1}{\tilde{\eta}-1}.$$

A2. In the framework of Theorem 3.1, for any fixed n, we put $J \triangleq [(\hat{y}+1)n]$, where $[\cdot]$ takes the integer part, and

$$w^{\varphi^*}(i) \triangleq \begin{cases} v^{\varphi^*}(i/n), & \text{if } i \leq J; \\ 0, & \text{if } i > J. \end{cases} \tag{3.19}$$

By applying the Dynkin's formula (see (Bremaud 1999, p.378-380, Thm.2.2)) to the bounded function w^{φ^*}, we have, for any control policy Φ,

$$\begin{aligned} & W^\Phi(i) \\ = \; & E_i^\Phi\left[\int_0^\infty c(\frac{^nY_s}{n},\Phi(^nY_s))ds\right] \end{aligned}$$

$$\begin{aligned} = \; & w^{\varphi^*}(i)+\lim_{t\to\infty}E_i^\Phi\left[\int_0^t\left\{c(\frac{^nY_s}{n},\Phi(^nY_s)) \; +n\lambda(\frac{^nY_s}{n},\Phi(^nY_s))w^{\varphi^*}(^nY_s+1)\right.\right. \\ & +n\mu(\frac{^nY_s}{n},\Phi(^nY_s))w^{\varphi^*}(^nY_s-1) \\ & \left.\left.-n(\lambda(\frac{^nY_s}{n},\Phi(^nY_s))+\mu(\frac{^nY_s}{n},\Phi(^nY_s))w^{\varphi^*}(^nY_s)\right\}ds\right]. \end{aligned} \tag{3.20}$$

Remember, $\lim_{t\to\infty}E_i^\Phi[w^{\varphi^*}(^nY_t)]=0$ because $\lim_{t\to\infty}{}^nY_t=0$ *a.s.*; more rigorous reasonings can be found in (Piunovskiy 2009b, Lem.2).

A3. For any fixed control policy Φ,

(a)

$$E_i^\Phi\left[\int_0^\infty I\{^nY_s=k\}n[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))+\mu(\frac{^nY_s}{n},\Phi(^nY_s))]ds\right]$$

$$\le \left\{\begin{array}{ll} \frac{\tilde\eta+1}{\tilde\eta-1}\times\frac{\tilde\eta^i-1}{\tilde\eta^k}, & \text{if } i\le k; \\ \\ \frac{\tilde\eta+1}{\tilde\eta-1}\times\frac{\tilde\eta^k-1}{\tilde\eta^k}, & \text{if } i\ge k \end{array}\right\} < \frac{\tilde\eta+1}{\tilde\eta-1};$$

(b)

$$E_i^\Phi\left[\int_0^\infty I\{^nY_s\le k\}n[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))+\mu(\frac{^nY_s}{n},\Phi(^nY_s))]ds\right]\le\frac{(\tilde\eta+1)k}{\tilde\eta-1}.$$

For the proofs, see (Piunovskiy 2009b, Lem.5, Rem.4), to which very similar reasonings result in the following:

(c) For any $k=1,2,\ldots,$

$$E_i^\Phi\left[\int_0^\infty \eta_1^{^nY_s/n}I\{^nY_s\ge k\}n[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))+\mu(\frac{^nY_s}{n},\Phi(^nY_s))]ds\right]\le\tilde V(i),$$

where

$$\tilde V(i)=\left\{\begin{array}{ll} \frac{(\tilde\eta+1)\eta_1^{(i+1)/n}}{(\tilde\eta-\eta_1^{1/n})(\eta_1^{1/n}-1)}+Z_1, & \text{if } i>k; \\ \\ Z_2(\tilde\eta^i-1), & \text{if } i\le k \end{array}\right.$$

satisfies $\sup_{i>0} \frac{|\tilde{V}(i)|}{\eta_1^{i/n}} < \infty$ and solves

$$\eta_1^{i/n} I\{i \geq k\} + \frac{1}{1+\tilde{\eta}} \tilde{V}(i+1) + \frac{\tilde{\eta}}{1+\tilde{\eta}} \tilde{V}(i-1) - \tilde{V}(i) = 0,$$

$$\tilde{V}(0) = 0.$$

Here,

$$Z_1 \triangleq \frac{\eta_1^{k/n}(1+\tilde{\eta})[\eta_1^{1/n}(\tilde{\eta}^{k-1}-1) - (\tilde{\eta}^k - 1)]}{(\tilde{\eta} - \eta_1^{1/n})(\tilde{\eta}-1)\tilde{\eta}^{k-1}(\eta_1^{1/n}-1)},$$

$$Z_2 \triangleq \frac{\eta_1^{k/n}(\tilde{\eta}+1)}{(\tilde{\eta} - \eta_1^{1/n})(\tilde{\eta}-1)\tilde{\eta}^{k-1}}.$$

The more detailed calculations for $\tilde{V}$ are given at the end of this section.

Proof of Theorem 3.1. (a) In what follows, n is fixed, $y_1, y_2, \ldots$ are from Definition 3.1 corresponding to φ^*, and $L \triangleq \max\{j:\ y_j < \hat{y}+1\}$. Recall that $J \triangleq [(\hat{y}+1)n]$, where $[\cdot]$ takes the integer part, and $w^{\varphi^*}(i)$ is given by (3.19). Let us put

$$\bar{S} \triangleq \{j \in S:\ j < J;\ \forall l = 1,2,\ldots,L, y_l \notin [(j-1)/n,\ (j+1)/n]\}.$$

According to (3.20), for any control policy Φ, for any fixed $i \leq \hat{y}n$,

$${}^nW^{\Phi}(i) = w^{\varphi^*}(i) + \mathscr{E}_1^{\Phi} + \mathscr{E}_2^{\Phi} + \mathscr{E}_3^{\Phi},$$

where

$$\begin{aligned}
\mathscr{E}_1^{\Phi} &= E_i^{\Phi}\Bigg[\int_0^{\infty} \{c(\frac{{}^nY_s}{n}, \Phi({}^nY_s)) + n\lambda(\frac{{}^nY_s}{n}, \Phi({}^nY_s)) w^{\varphi^*}({}^nY_s+1) \\
&\quad + n\mu(\frac{{}^nY_s}{n}, \Phi({}^nY_s)) w^{\varphi^*}({}^nY_s-1) \\
&\quad - n[\lambda(\frac{{}^nY_s}{n}, \Phi({}^nY_s)) + \mu(\frac{{}^nY_s}{n}, \Phi({}^nY_s))] w^{\varphi^*}({}^nY_s)\} I\{{}^nY_s \in \bar{S}\} ds\Bigg],
\end{aligned}$$

$$\begin{aligned}
\mathscr{E}_2^{\Phi} &= E_i^{\Phi}\Bigg[\int_0^{\infty} \{c(\frac{{}^nY_s}{n}, \Phi({}^nY_s)) + n\lambda(\frac{{}^nY_s}{n}, \Phi({}^nY_s)) w^{\varphi^*}({}^nY_s+1) \\
&\quad + n\mu(\frac{{}^nY_s}{n}, \Phi({}^nY_s)) w^{\varphi^*}({}^nY_s-1) \\
&\quad - n[\lambda(\frac{{}^nY_s}{n}, \Phi({}^nY_s)) + \mu(\frac{{}^nY_s}{n}, \Phi({}^nY_s))] w^{\varphi^*}({}^nY_s)\} I\{{}^nY_s \notin \bar{S}\} I\{{}^nY_s < J\} ds\Bigg],
\end{aligned}$$

and

$$
\begin{aligned}
\mathscr{E}_3^{\Phi} &= E_i^{\Phi}\Big[\int_0^{\infty}\{c(\frac{^nY_s}{n},\Phi(^nY_s))+n\lambda(\frac{^nY_s}{n},\Phi(^nY_s))w^{\varphi^*}(^nY_s+1) \\
&\quad +n\mu(\frac{^nY_s}{n},\Phi(^nY_s))w^{\varphi^*}(^nY_s-1) \\
&\quad -n[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))+\mu(\frac{^nY_s}{n},\Phi(^nY_s))]w^{\varphi^*}(^nY_s)\}I\{^nY_s\geq J\}ds\Big].
\end{aligned}
$$

Now by applying the mean value theorem, we have

$$
\begin{aligned}
\mathscr{E}_1^{\Phi} &= E_i^{\Phi}\Bigg[\int_0^{\infty}\Bigg\{c(\frac{^nY_s}{n},\Phi(^nY_s))+\lambda(\frac{^nY_s}{n},\Phi(^nY_s))\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y={^nY_s}/n} \\
&\quad -\mu(\frac{^nY_s}{n},\Phi(^nY_s))\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y={^nY_s}/n} \\
&\quad +\lambda(\frac{^nY_s}{n},\Phi(^nY_s))\left[\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y=\theta_1\in({^nY_s}/n,({^nY_s}+1)/n)}-\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y={^nY_s}/n}\right] \\
&\quad -\mu(\frac{^nY_s}{n},\Phi(^nY_s))\left[\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y=\theta_2\in(({^nY_s}-1)/n,{^nY_s}/n)}-\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y={^nY_s}/n}\right]\Bigg\} \\
&\quad \times I\{^nY_s\in\bar{S}\}ds\Bigg],
\end{aligned}
\tag{3.21}
$$

so that

$$
\begin{aligned}
\mathscr{E}_1^{\Phi} &= E_i^{\Phi}\Bigg[\int_0^{\infty}\Bigg\{c(\frac{^nY_s}{n},\Phi(^nY_s))+\left[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))-\mu(\frac{^nY_s}{n},\Phi(^nY_s))\right]\frac{dv}{dy}\bigg|_{y=\frac{^nY_s}{n}} \\
&\quad +\left[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))-\mu(\frac{^nY_s}{n},\Phi(^nY_s))\right]\left[\frac{dv^{\varphi^*}}{dy}\bigg|_{y=\frac{^nY_s}{n}}-\frac{dv}{dy}\bigg|_{y=\frac{^nY_s}{n}}\right]\Bigg\} \\
&\quad \times I\{^nY_s\in\bar{S}\}ds\Bigg]+\gamma^{\Phi},
\end{aligned}
$$

where

$$
\begin{aligned}
\gamma^{\Phi} &= E_i^{\Phi}\Bigg[\int_0^{\infty}\Bigg\{\lambda(\frac{^nY_s}{n},\Phi(^nY_s))\left[\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y=\theta_1\in(\frac{^nY_s}{n},\frac{^nY_s+1}{n})}-\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y=\frac{^nY_s}{n}}\right] \\
&\quad -\mu(\frac{^nY_s}{n},\Phi(^nY_s))\left[\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y=\theta_2\in(\frac{^nY_s-1}{n},\frac{^nY_s}{n})}-\frac{dv^{\varphi^*}(y)}{dy}\bigg|_{y=\frac{^nY_s}{n}}\right]\Bigg\} \\
&\quad \times I\{^nY_s\in\bar{S}\}ds\Bigg],
\end{aligned}
$$

and according to **A3**(b),

$$\begin{aligned}|\gamma^{\Phi}| &\le \frac{D}{n^2}E_i^{\Phi}\left[\int_0^{\infty} n[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))+\mu(\frac{^nY_s}{n},\Phi(^nY_s))]I\{^nY_s\le J\}ds\right]\\ &\le \frac{D(\hat{y}+1)}{n}\times\frac{\tilde{\eta}+1}{\tilde{\eta}-1}.\end{aligned}$$

Here, we recall that D is the common Lipschitz constant of $\frac{dv^{\varphi^*}}{dy}=\frac{c(y,\varphi^*(y))}{\mu(y,\varphi^*(y))-\lambda(y,\varphi^*(y))}$ on all of the intervals (y_j,y_{j+1}), $j=0,1,\ldots,L$.

Therefore, according to (3.17) and **A3**(b), for an arbitrary control policy Φ,

$$\begin{aligned}\mathscr{E}_1^{\Phi} &\ge \frac{-1}{n}\sup_{^nY_s\in\tilde{S}}\left[\left.\frac{dv^{\varphi^*}}{dy}\right|_{y=\frac{^nY_s}{n}}-\left.\frac{dv}{dy}\right|_{y=\frac{^nY_s}{n}}\right]\\ &\quad\times E_i^{\Phi}\left[\int_0^{\infty} n[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))+\mu(\frac{^nY_s}{n},\Phi(^nY_s))]I\{^nY_s\le J\}ds\right]-|\gamma^{\Phi}|\\ &\ge -\sup_{^nY_s\in\tilde{S}}\left[\left.\frac{dv^{\varphi^*}}{dy}\right|_{y=\frac{^nY_s}{n}}-\left.\frac{dv}{dy}\right|_{y=\frac{^nY_s}{n}}\right]\frac{(\hat{y}+1)(\tilde{\eta}+1)}{\tilde{\eta}-1}\\ &\quad-\frac{D(\hat{y}+1)}{n}\times\frac{\tilde{\eta}+1}{\tilde{\eta}-1}.\end{aligned} \tag{3.22}$$

On the other hand, for $\Phi=\Phi^*$, it follows from (3.4) and (3.21) that $\mathscr{E}_1^{\Phi^*}=\gamma^{\Phi^*}$, which leads to

$$|\mathscr{E}_1^{\Phi^*}|=|\gamma^{\Phi^*}|\le\frac{D(\hat{y}+1)}{n}\times\frac{\tilde{\eta}+1}{\tilde{\eta}-1}. \tag{3.23}$$

We emphasize that (3.23) does not follow directly from (3.22). In fact, (3.22) is only needed for proving part (b), and we just prove it here incidentally.

For $\mathscr{E}_2^{\Phi}$, we estimate

$$\begin{aligned}|\mathscr{E}_2^{\Phi}| &\le E_i^{\Phi}\left[\int_0^{\infty}\left\{\frac{|c(\frac{^nY_s}{n},\Phi(^nY_s))|}{n[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))+\mu(\frac{^nY_s}{n},\Phi(^nY_s))]}\right.\right.\\ &\quad\left.+\frac{C_1\eta_1^{\frac{^nY_s}{n}}(\eta_1^{1/n}-1)}{\ln\eta_1}\times\frac{\tilde{\eta}+1}{\tilde{\eta}-1}\right\}\\ &\quad\left.\times\, n[\lambda(\frac{^nY_s}{n},\Phi(^nY_s))+\mu(\frac{^nY_s}{n},\Phi(^nY_s))]\sum_{j<J,j\notin\tilde{S}}I\{^nY_s=j\}ds\right].\end{aligned}$$

Here and below, we use (3.19), Conditions 3.1 and Statements **A1** and **A3**(a). Thus

$$|\mathscr{E}_2^{\Phi}|\le\left\{\frac{C_1\eta_1^{\hat{y}+1}}{n}+\frac{C_1\eta_1^{\hat{y}+1}(\eta_1^{1/n}-1)}{\ln\eta_1}\times\frac{\tilde{\eta}+1}{\tilde{\eta}-1}\right\}\frac{3L(\tilde{\eta}+1)}{\tilde{\eta}-1}. \tag{3.24}$$

Recall that the set $\{j<J:\ j\notin\tilde{S}\}$ contains no more than $3L$ points.

For $\mathscr{E}_3^{\Phi}$, we know from **A1** that for all i,

$$|w^{\varphi^*}(i)| \leq C_1 \times \frac{\eta_1^{J/n}}{\ln \eta_1} \times \frac{\tilde{\eta}+1}{\tilde{\eta}-1}.$$

Here, we recall that $w^{\varphi^*}(i) = 0$ if $i > J$, see (3.19). Therefore, using Conditions 3.1, we obtain

$$\begin{aligned}
|\mathscr{E}_3^{\Phi}| &\leq E_i^{\Phi}\left[\int_0^{\infty}\left\{\frac{|c(\frac{{}^nY_s}{n},\Phi({}^nY_s))|}{n[\lambda(\frac{{}^nY_s}{n},\Phi({}^nY_s))+\mu(\frac{{}^nY_s}{n},\Phi({}^nY_s))]}+2C_1\times\frac{\eta_1^{J/n}}{\ln\eta_1}\times\frac{\tilde{\eta}+1}{\tilde{\eta}-1}\right\}\right.\\
&\quad\left.\times\, n[\lambda(\frac{{}^nY_s}{n},\Phi({}^nY_s))+\mu(\frac{{}^nY_s}{n},\Phi({}^nY_s))]I\{{}^nY_s\geq J\}ds\right]\\
&\leq C_1\left(1+\frac{2}{\ln\eta_1}\times\frac{\tilde{\eta}+1}{\tilde{\eta}-1}\right)E_i^{\Phi}\left[\int_0^{\infty}I\{{}^nY_s\geq J\}\eta_1^{\frac{{}^nY_s}{n}}\right.\\
&\quad\left.\times\, n[\lambda(\frac{{}^nY_s}{n},\Phi({}^nY_s))+\mu(\frac{{}^nY_s}{n},\Phi({}^nY_s))]ds\right].
\end{aligned}$$

Finally, since $i \leq \hat{y}n \leq (\hat{y}+1)n - 1 \leq J$, according to **A3**(c), we have

$$|\mathscr{E}_3^{\Phi}| \leq C_1\left(1+\frac{2}{\ln\eta_1}\times\frac{\tilde{\eta}+1}{\tilde{\eta}-1}\right)\times\frac{\eta_1^{\hat{y}+1}(\tilde{\eta}+1)\tilde{\eta}^2}{(\tilde{\eta}-\eta_1)(\tilde{\eta}-1)}\tilde{\eta}^{-n}. \tag{3.25}$$

Now inequality (3.13) follows directly from (3.23), (3.24) and (3.25). Thus, part (a) is proved.

(b) Bearing in mind inequality (3.22), we obtain

$$\begin{aligned}
\inf_{\Phi} {}^nW^{\Phi}(i) &\geq v^{\varphi^*}(i/n) - \sup_{y\in(0,\hat{y}+1]}\left[\frac{dv^{\varphi^*}}{dy}-\frac{dv}{dy}\right]\frac{(\hat{y}+1)(\tilde{\eta}+1)}{\tilde{\eta}-1}\\
&\quad -\frac{K_1}{n}-\frac{K_2}{\tilde{\eta}^n}-K_3(\eta_1^{1/n}-1),
\end{aligned}$$

and inequality (3.14) follows. Thus, part (b) is proved. ■

Proof of Corollary 3.1. This corollary trivially holds. ■

Calculations involved in Example 3.2. To study the fluid model, we shall solve the Bellman equation

$$\inf_{a\in[0,1]}\left\{\frac{dv}{dz}(d_0+d_1a-1)+z-Ra\right\}=0,\ v(0)=0.$$

One can verify that the solution to the fluid model is :

$$\varphi^*(y)=\begin{cases}1, \ \forall 0<y\leq y^* \triangleq \frac{(1-d_0)R}{d_1};\\ 0, \ \forall y>y^*,\end{cases}$$

and

$$v(y) = \begin{cases} \frac{Ry - y^2/2}{d_0+d_1-1}, \ \forall 0 < y \le y^*; \\ \frac{y^2}{(1-d_0)2} + \frac{Ry^* - \frac{y^{*2}}{2}}{d_0+d_1-1} - \frac{y^{*2}}{2(1-d_0)}, \ \forall y > y^*. \end{cases}$$

Back to the stochastic problem (3.1), we shall solve the Bellman equation

$$\begin{aligned} &\inf_{a \in [0,1]} \left\{ \frac{i}{n} - Ra + (d_0 + d_1 a) n \, {}^nV(i+1) + n \, {}^nV(i-1) \right. \\ &\left. -n(d_0 + d_1 a + 1) \, {}^nV(i) \right\} = 0, \\ &{}^nV(0) = 0. \end{aligned}$$

The preliminary calculations are given as below. Then we verify the solutions obtained in that way truly solve the Bellman equation written above.

Firstly, we make the following guess: when $i > i^*$, the optimal policy is given by $\Phi^*(i) = 1$, and when $i \le i^*$, the optimal policy is given by $\Phi^*(i) = 0$. Under this guess we start off with finding ${}^nV(i)$ in terms of the unknown yet fixed value i^*.

Consider the case $i \le i^*$. Then the Bellman function should satisfy the equation:

$$\begin{aligned} &\frac{i}{n} - R + (d_0 + d_1) n \, {}^nV(i+1) + n \, {}^nV(i-1) - (d_0 + d_1 + 1) n \, {}^nV(i) = 0, \\ &{}^nV(0) = 0. \end{aligned}$$

To solve this equation, after observing the linear form with respect to ${}^nV(\cdot)$ on the left hand side, we first consider the homogeneous case:

$$(d_0 + d_1) n \, {}^nV(i+1) + n \, {}^nV(i-1) - (d_0 + d_1 + 1) n \, {}^nV(i) = 0.$$

We guess the solution in the homogeneous case is in the form ${}^nV^c(i) = D_2 \alpha^i + D_1$. Then by substitution we have

$$\begin{aligned} &(d_0 + d_1) n (D_2 \alpha^{i+1} + D_1) + n (D_2 \alpha^{i-1} + D_1) - (d_0 + d_1 + 1) n (D_2 \alpha^i + D_1) = 0 \\ \Leftrightarrow \quad &(d_0 + d_1) \alpha^2 - (d_0 + d_1 + 1) \alpha + 1 = 0 \\ \Leftrightarrow \quad &\alpha_1 = 1, \ \alpha_2 = \frac{1}{d_0 + d_1} \Rightarrow {}^nV^c(i) = D_1 + D_2 \left(\frac{1}{d_0 + d_1} \right)^i, \end{aligned}$$

where in the process we have ignored the case $D_2 = 0$.

Then we consider the inhomogeneous case:

$$\frac{i}{n} - R + (d_0 + d_1) n \, {}^nV(i+1) + n \, {}^nV(i) - (d_0 + d_1 + 1) n \, {}^nV(i) = 0,$$

for which we guess the solution is in the form ${}^{n}V^{p}(i) = \frac{Ai^2+Bi}{n^2}$. By substitution we have

$$
\begin{aligned}
& \frac{i}{n} - R + (d_0+d_1)\frac{A(i+1)^2+B(i+1)}{n} + \frac{A(i-1)^2+B(i-1)}{n} \\
& -(d_0+d_1+1)\frac{Ai^2+Bi}{n} = 0 \\
\Leftrightarrow \quad & i(2A(d_0+d_1)+1-2A)+((d_0+d_1)(A+B)+A-B-nR) = 0
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \quad & A = \frac{1}{2(1-d_0-d_1)}, B = \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2} \\
\Rightarrow \quad & {}^{n}V^{p}(i) = \frac{\frac{i^2}{2(1-d_0-d_1)} + \left(\frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right)i}{n^2}.
\end{aligned}
$$

Therefore, the general solution is in the form

$$
\begin{aligned}
{}^{n}V(i) &= {}^{n}V^{p}(i) + {}^{n}V^{c}(i) = D_1 + D_2\left(\frac{1}{d_0+d_1}\right)^i \\
&+ \frac{\frac{i^2}{2(1-d_0-d_1)} + \left(\frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right)i}{n^2}.
\end{aligned}
$$

With in mind the condition ${}^{n}V(0) = 0$ we further have $D_1 = -D_2$, and thus finally have

$$
\begin{aligned}
{}^{n}V(i) &= {}^{n}V^{p}(i) + {}^{n}V^{c}(i) \\
&= D_1 - D_1\left(\frac{1}{d_0+d_1}\right)^i \\
&+ \frac{\frac{i^2}{2(1-d_0-d_1)} + \left(\frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right)i}{n^2}.
\end{aligned} \tag{3.26}
$$

Now consider the case $i > i^*$, where the Bellman function should satisfy

$$\frac{i}{n} + d_0 n\,{}^{n}V(i+1) + n\,{}^{n}V(i-1) - (d_0+1)n\,{}^{n}V(i) = 0.$$

Similarly as in the case $i \leq i^*$, we consider the homogeneous case and inhomogeneous case, in order to solve for ${}^{n}V(i)$.

In the homogeneous case $d_0 n\,{}^{n}V(i+1) + n\,{}^{n}V(i-1) - (d_0+1)n\,{}^{n}V(i) = 0$, we guess the solution should be in the form ${}^{n}V^{c}(i) = D_4\alpha^i + D_3$, and by substitution and excluding the case $D_4 = 0$ we have

$$
\begin{aligned}
& d_0\alpha^2 + 1 - (d_0+1)\alpha = 0 \Leftrightarrow \alpha_1 = 1,\ \alpha_2 = 1/d_0 \\
\Rightarrow \quad & {}^{n}V^{c}(i) = D_3 + D_4(1/d_0)^i.
\end{aligned}
$$

In the inhomogeneous case $\frac{i}{n} + d_0 n\,{}^{n}V(i+1) + n\,{}^{n}V(i-1) - (d_0+1)n\,{}^{n}V(i) = 0$, we make the guess ${}^{n}V^{p}(i) = \frac{Ai^2+Bi}{n^2}$ and by substitution we have

$$
\begin{aligned}
& i(1+2A(d_0-1))+A(d_0+1)+B(d_0-1)=0 \\
\Leftrightarrow \quad & A=\frac{1}{2(1-d_0)},\ B=\frac{1+d_0}{2(1-d_0)^2} \Rightarrow {}^{n}V^{p}(i)=\frac{\frac{1}{2(1-d_0)}i^2+\frac{1+d_0}{2(1-d_0)^2}i}{n^2}.
\end{aligned}
$$

Eventually, the general solution is provided by

$$
{}^{n}V(i)=\frac{\frac{1}{2(1-d_0)}i^2+\frac{1+d_0}{2(1-d_0)^2}i}{n^2}+D_3+D_4(1/d_0)^i, \tag{3.27}
$$

where we must ignore the rapidly growing solution meaning $D_4=0$. (See (Piunovskiy 2009b, Lem.2,4.5).)

So as to determine D_1 and D_3, let us consider two particular values of i.

When $i=i^*$: by equating ${}^{n}V(i^*)$ given by (3.26) to that given by (3.27) we have

$$
\begin{aligned}
& D_1-D_1\left(\frac{1}{d_0+d_1}\right)^{i^*}+\frac{\frac{i^{*2}}{2(1-d_0-d_1)}+\left(\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right)i^*}{n^2} \\
= \quad & D_3+\frac{\frac{i^{*2}}{2(1-d_0)}+\frac{(1+d_0)i^*}{2(1-d_0)^2}}{n^2}
\end{aligned} \tag{3.28}
$$

When $i=i^*+1$: by equating ${}^{n}V(i^*+1)$ given by (3.26) to that given by (3.27) we have

$$
\begin{aligned}
& D_1-D_1\left(\frac{1}{d_0+d_1}\right)^{i^*+1}+\frac{\frac{(i^*+1)^2}{2(1-d_0-d_1)}+\left(\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right)(i^*+1)}{n^2} \\
= \quad & D_3+\frac{\frac{(i^*+1)^2}{2(1-d_0)}+\frac{(1+d_0)(i^*+1)}{2(1-d_0)^2}}{n^2}
\end{aligned} \tag{3.29}
$$

Solve equations (3.28) and (3.29) together and we obtain:

$$
\begin{aligned}
& D_1=\frac{\frac{2i^*+1}{2(1-d_0-d_1)}+\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}-\frac{2i^*+1}{2(1-d_0)}-\frac{1+d_0}{2(1-d_0)^2}}{n^2\left(\frac{1}{(d_0+d_1)^{i^*}}\left(\frac{1}{d_0+d_1}-1\right)\right)} \\
\Rightarrow \quad & D_3=\left(1-\frac{1}{(d_0+d_1)^{i^*}}\right) \\
& \times\left\{\frac{\frac{2i^*+1}{2(1-d_0-d_1)}+\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}-\frac{2i^*+1}{2(1-d_0)}-\frac{1+d_0}{2(1-d_0)^2}}{n^2\left(\frac{1}{(d_0+d_1)^{i^*}}\left(\frac{1}{d_0+d_1}-1\right)\right)}\right\} \\
& +\frac{\frac{i^{*2}}{2(1-d_0-d_1)}+\left(\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right)i^*}{n^2}-\frac{\frac{i^{*2}}{2(1-d_0)}+\frac{(1+d_0)i^*}{2(1-d_0)^2}}{n^2}.
\end{aligned}
$$

Now one must determine the unknown i^*. To this end, let us consider

$$f(y,a) \triangleq \frac{y}{n} - Ra + (d_0 + d_1 a) n\, {}^{n}V(y+1) + n\, {}^{n}V(y-1) - n(d_0 + d_1 a + 1)\, {}^{n}V(y)$$

and

$$f_1(y) \triangleq \frac{\partial f}{\partial a} = -R + d_1 n({}^{n}V(y+1) - {}^{n}V(y)).$$

Under our guess, i^* must be such that

$$\frac{\partial f}{\partial a}(y^*, a) = -R + d_1 n({}^{n}V(y^*+1) - {}^{n}V(y^*)) = 0$$

where y^* is some number that satisfies $i^* = \lfloor y^* \rfloor$. Put ${}^{n}V(y^*+1), {}^{n}V(y^*)$ given by equation (3.27) into this equation and we obtain

$$\begin{aligned} \frac{nR}{d_1} &= \frac{2y^*+1}{2(1-d_0)} + \frac{1+d_0}{2(1-d_0)^2} \\ \Leftrightarrow \quad y^* &= \frac{nR(1-d_0)}{d_1} - \frac{1+d_0}{2(1-d_0)} - \frac{1}{2}. \end{aligned}$$

The Bellman equation is solved by the feedback policy

$$\Phi^*(i) = \begin{cases} 1, \ \forall 0 < i \le i^* \triangleq \left\lfloor \frac{nR(1-d_0)}{d_1} - \frac{1+d_0}{(1-d_0)2} - \frac{1}{2} \right\rfloor ; \\ 0, \ \forall i > i^*, \end{cases}$$

and the function

$$\begin{aligned} {}^{n}V(i) \;&=\; \frac{\frac{2i^*+1}{2(1-d_0-d_1)} + \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2} - \frac{2i^*+1}{2(1-d_0)} - \frac{1+d_0}{2(1-d_0)^2}}{n^2\left(\frac{1}{(d_0+d_1)^{i^*}}\left(\frac{1}{d_0+d_1} - 1\right)\right)} \\ &\times \left(1 - \left(\frac{1}{d_0+d_1}\right)^i\right) + \frac{\frac{i^2}{2(1-d_0-d_1)} + \left(\frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right) i}{n^2} \end{aligned}$$

when $i \le i^*$, and

$$\begin{aligned} {}^{n}V(i) \;&=\; \left(1 - \frac{1}{(d_0+d_1)^{i^*}}\right) \\ &\times \left\{ \frac{\frac{2i^*+1}{2(1-d_0-d_1)} + \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2} - \frac{2i^*+1}{2(1-d_0)} - \frac{1+d_0}{2(1-d_0)^2}}{n^2\left(\frac{1}{(d_0+d_1)^{i^*}}\left(\frac{1}{d_0+d_1} - 1\right)\right)} \right\} \\ &+ \frac{\frac{1}{2(1-d_0)} i^2 + \frac{1+d_0}{2(1-d_0)^2} i}{n^2} + \frac{\frac{i^{*2}}{2(1-d_0-d_1)} + \left(\frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right) i^*}{n^2} \\ &- \frac{\frac{i^{*2}}{2(1-d_0)} + \frac{(1+d_0)i^*}{2(1-d_0)^2}}{n^2} \end{aligned}$$

when $i > i^*$. To verify the optimality of the derived policy, let us focus on the negativity (positivity) of $f_1(i)$. Since for $''V(\cdot)$ given by (3.27) $''V(y+1) -{}'' V(y)$ increases when y increases, clearly $f_1(i) > 0\ \forall i > i^*$, we only verify the case when $i \leq i^*$ and $i^* > 0$ as follows.

By substituting $''V(y)$, $''V(y+1)$ given by (3.26) into $f_1(y)$ we have

$$
\begin{aligned}
f_1(y) &= -R + d_1 n \left\{ -D_1 \left(\frac{1}{d_0+d_1}\right)^{y+1} \right. \\
&\quad + \frac{\frac{(y+1)^2}{2(1-d_0-d_1)} + \left[\frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right](y+1)}{n^2} \\
&\quad \left. + D_1 \left(\frac{1}{d_0+d_1}\right)^{y} - \frac{\frac{y^2}{2(1-d_0-d_1)} + \left[\frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right] y}{n^2} \right\} \\
&= -R + d_1 n \left\{ D_1 \left(\frac{1}{d_0+d_1}\right)^{y} \left(1 - \frac{1}{d_0+d_1}\right) \right. \\
&\quad \left. + \frac{\frac{2y+1}{2(1-d_0-d_1)} + \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}}{n^2} \right\} \\
&= -R + d_1 n \left\{ \frac{\frac{2i^*+1}{2(1-d_0-d_1)} + \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2} - \frac{2i^*+1}{2(1-d_0)} - \frac{1+d_0}{2(1-d_0)^2}}{n^2 \left(\frac{1}{(d_0+d_1)^{i^*}} \left(\frac{1}{d_0+d_1} - 1\right)\right)} \right. \\
&\quad \times \left(\frac{1}{d_0+d_1}\right)^{y} \left(1 - \frac{1}{d_0+d_1}\right) \\
&\quad \left. + \frac{\frac{2y+1}{2(1-d_0-d_1)} + \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}}{n^2} \right\} \\
&= -R + d_1 \left\{ \frac{\frac{2i^*+1}{2(1-d_0-d_1)} + \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2} - \frac{2i^*+1}{2(1-d_0)} - \frac{1+d_0}{2(1-d_0)^2}}{n \left(\frac{1}{(d_0+d_1)^{i^*}} \left(\frac{1}{d_0+d_1} - 1\right)\right)} \right. \\
&\quad \times \left(\frac{1}{d_0+d_1}\right)^{y} \left(1 - \frac{1}{d_0+d_1}\right) \\
&\quad \left. + \frac{\frac{2y+1}{2(1-d_0-d_1)} + \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2}}{n} \right\} \\
&= -d_1 \left\{ \frac{\frac{2i^*+1}{2(1-d_0-d_1)} + \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2} - \frac{2i^*+1}{2(1-d_0)} - \frac{1+d_0}{2(1-d_0)^2}}{n \left(\frac{1}{(d_0+d_1)^{i^*}}\right)} \right. \\
&\quad \left. \times \left(\frac{1}{d_0+d_1}\right)^{y} \right\} + d_1 \frac{\frac{2y+1}{2(1-d_0-d_1)} + \frac{nR}{d_0+d_1-1} + \frac{d_0+d_1+1}{2(1-d_0-d_1)^2} - \frac{nR}{d_1}}{n}
\end{aligned}
$$

$$
\begin{aligned}
&= -d_1\left\{\frac{\frac{2i^*+1}{2(1-d_0-d_1)}+\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}-\frac{2i^*+1}{2(1-d_0)}-\frac{1+d_0}{2(1-d_0)^2}}{n\left(\frac{1}{(d_0+d_1)^{i^*}}\right)}\right. \\
&\quad \left.\times\left(\frac{1}{d_0+d_1}\right)^y\right\} \\
&\quad +d_1\frac{\frac{2y+1}{2(1-d_0-d_1)}+\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}-\frac{2y^*+1}{2(1-d_0)}-\frac{1+d_0}{2(1-d_0)^2}}{n}.
\end{aligned}
$$

Therefore $f_1(i^*) \leq 0$.

Furthermore,

$$
\begin{aligned}
\frac{d(nf_1(y))}{dy} &= \frac{2d_1}{2(1-d_0-d_1)}-d_1\left[\frac{2i^*+1}{2(1-d_0-d_1)}+\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right. \\
&\quad \left.-\frac{2i^*+1}{2(1-d_0)}-\frac{1+d_0}{2(1-d_0)^2}\right]\times\frac{(d_0+d_1)^{i^*}}{(d_0+d_1)^y}\ln\frac{1}{d_0+d_1} \\
&\geq d_1\left\{\frac{1}{1-d_0-d_1}-\left[\frac{2i^*+1}{2(1-d_0-d_1)}+\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right.\right. \\
&\quad \left.\left.-\frac{2i^*+1}{2(1-d_0)}-\frac{1+d_0}{2(1-d_0)^2}\right]\ln\frac{1}{d_0+d_1}\right\},
\end{aligned}
$$

where we have assumed that the term in the square bracket is positive, and the case of a negative term in the square bracket is trivial. Let us estimate the term in the parenthesis as follows:

$$
\begin{aligned}
&\frac{1}{1-d_0-d_1}-\left[\frac{2i^*+1}{2(1-d_0-d_1)}+\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right. \\
&\quad \left.-\frac{2i^*+1}{2(1-d_0)}-\frac{1+d_0}{2(1-d_0)^2}\right]\ln\frac{1}{d_0+d_1} \\
\geq\ &\frac{1}{1-d_0-d_1}-\left[\frac{2i^*+1}{2(1-d_0-d_1)}+\frac{nR}{d_0+d_1-1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)^2}\right. \\
&\quad \left.-\frac{2i^*+1}{2(1-d_0)}-\frac{1+d_0}{2(1-d_0)^2}\right]\left(\frac{1-d_0-d_1}{d_0+d_1}\right) \\
\geq\ &\frac{1}{1-d_0-d_1}-\left[\frac{2y^*+1}{2(d_0+d_1)}-\frac{nR}{d_0+d_1}+\frac{d_0+d_1+1}{2(1-d_0-d_1)(d_1+d_0)}\right. \\
&\quad \left.-\frac{nR(1-d_0-d_1)}{d_1(d_0+d_1)}\right] \\
=\ &-\frac{2y^*+1}{2(d_0+d_1)}+\frac{nR(1-d_0)}{(d_0+d_1)d_1}+\frac{d_0+d_1-1}{2(1-d_0-d_1)(d_0+d_1)} \\
=\ &-\frac{2y^*+1}{2(d_0+d_1)}+\frac{2nR(1-d_0)}{2(d_0+d_1)d_1}-\frac{1}{2(d_0+d_1)} \\
=\ &\frac{-d_1}{2(d_0+d_1)d_1}-\frac{2nR(1-d_0)-\frac{(1+d_0)d_1}{(1-d_0)}}{2(d_0+d_1)d_1}+\frac{2nR-2nRd_0}{2(d_0+d_1)d_1}>0,
\end{aligned}
$$

meaning that $f_1(i) \le 0$ when $i \le i^*$, as required.

Finally, without many efforts one can also verify that the Bellman equation is satisfied by the solutions derived from the preliminarily derived $''V$. ■

Calculations involved in Example 3.3. To study the fluid model, let us solve the Bellman equation

$$\inf_{a \in [0.1]} \left\{ (d_0 + d_1 a - 1 - y)\frac{dv}{dy} + y + 1 - Ra \right\} = 0, v(0) = 0.$$

Firstly, we guess that for some values of R the optimal policy is provided by

$$\varphi^*(y) = \begin{cases} 0, \ \forall 0 < y \le y^*; \\ 1, \ \forall y > y^*, \end{cases}$$

where y^* is a constant yet unknown.

Under this guess, $\forall y \le y^*$ we have

$$\begin{aligned} & (d_1 - 1 - y)\frac{dv}{dy} + y + 1 = 0 \\ \Leftrightarrow \quad & \frac{dv}{dy} = \frac{y+1}{y+1-d_0} = 1 + \frac{d_0}{y+1-d_0} \\ \Leftrightarrow \quad & v(y) = y + d_0 \ln(y+1-d_0) + A, \end{aligned}$$

and the constant A is determined by the initial condition $v(0) = 0$ leading to

$$v(y) = y + d_0 \ln(y+1-d_0) - d_0 \ln(1-d_0). \tag{3.30}$$

$\forall y > y^*$ we have

$$\begin{aligned} & (d_0 + d_1 - 1 - y)\frac{dv}{dy} + y + 1 - R = 0 \\ \Leftrightarrow \quad & (d_0 + d_1 - y - 1)\frac{dv}{dy} = R - 1 - y \\ \Leftrightarrow \quad & \frac{dv}{dy} = 1 + \frac{d_0 + d_1 - R}{y + 1 - d_0 - d_1} \\ \Leftrightarrow \quad & v(y) = y + (d_0 + d_1 - R)\ln(y + 1 - d_0 - d_1) + D_5, \end{aligned} \tag{3.31}$$

where the local constant D_5 is determined by equating $v(y^*)$ given by (3.30) to the one given by (3.31):

$$
\begin{aligned}
& y^* + d_0 \ln(y^* + 1 - d_0) - d_0 \ln(1 - d_0) \\
& = y^* + (d_0 + d_1 - R)\ln(y^* + 1 - d_0 - d_1) + D_5 \\
\Leftrightarrow \quad & D_5 = d_0 \ln(y^* + 1 - d_0) - d_0 \ln(1 - d_0) - (d_0 + d_1 - R)\ln(y^* + 1 - d_0 - d_1) \\
\Leftrightarrow \quad & v(y) = y + (d_0 + d_1 - R)\ln(y + 1 - d_0 - d_1) + d_0 \ln(y^* + 1 - d_0) \\
& -d_0 \ln(1 - d_0) - (d_0 + d_1 - R)\ln(y^* + 1 - d_0 - d_1)
\end{aligned}
$$

when $y > y^*$.

Now to determine y^*, one should consider

$$
b(y,a) \triangleq (d_0 + d_1 a - 1 - y)\frac{dv}{dy} + y + 1 - Ra
$$

and

$$
b_1(y) \triangleq \frac{\partial b(y,a)}{\partial a} = d_1 \frac{dv}{dy} - R.
$$

Under our guess y^* must be such that $b_1(y^*) = 0$, meaning that from (3.30)

$$
\begin{aligned}
& d_1 \frac{y^* + 1}{y^* + 1 - d_0} - R = 0 \\
\Leftrightarrow \quad & y^* = \frac{R - Rd_0 - d_1}{d_1 - R}
\end{aligned}
$$

from which we see, to validate our guess, it is necessary to have $R \in (d_1, \frac{d_1}{1-d_0})$.

Therefore, when $d_1 < R < \frac{d_1}{1-d_0}$, the Bellman equation for the fluid model is solved by

$$
\varphi^*(y) = \begin{cases} 0, \ \forall 0 < y \leq y^*; \\ 1, \ \forall y > y^*, \end{cases}
$$

and

$$
v(y) \quad = \quad y + d_0 \ln(y + 1 - d_0) - d_0 \ln(1 - d_0),
$$

when $0 < y \leq y^*$ and

$$
\begin{aligned}
v(y) \quad = \quad & y + (d_0 + d_1 - R)\ln(y + 1 - d_0 - d_1) + d_0 \ln(y^* + 1 - d_0) - d_0 \ln(1 - d_0) \\
& -(d_0 + d_1 - R)\ln(y^* + 1 - d_0 - d_1),
\end{aligned}
$$

when $y > y^*$.

In case $0 < R \leq d_1$ (resp. $R \geq \frac{d_1}{1-d_0}$), the optimal policy is $\varphi^*(y) = 0$ (resp. $\varphi^*(y) = 1$) identically, and $v(y)$ is given by (3.30) (resp. (3.31)).

Finally let us verify the optimality of the aforementioned policies for the three sets of values of R, respectively. That is, we shall examine the positivity (negativity) of the function $b_1(y)$. Clearly, $b_1(y^*) = 0$ from the above calculations.

When $R \in (d_1, \frac{d_1}{1-d_0})$ and $y < y^*$:

$$b_1(y) \quad = \quad d_1 \frac{y+1}{y+1-d_0} - R > \frac{d_1}{1-d_0} - R > 0,$$

as required.

When $R \in (d_1, \frac{d_1}{1-d_0})$ and $y > y^*$

$$b_1(y) \quad = \quad \left(\frac{d_0 + d_1 - R}{y+1-d_0-d_1} + 1 \right) d_1 - R < 0,$$

as required.

The cases of $R \leq d_1$ and $R \geq \frac{d_1}{1-d_0}$ can be verified in the same way. ■

Calculations involved in A3(c) of Section 3.7. Let us solve

$$\eta_1^{i/n} I\{i \geq k\} + \frac{1}{1+\tilde{\eta}} \tilde{V}(i+1) + \frac{\tilde{\eta}}{1+\tilde{\eta}} \tilde{V}(i-1) - \tilde{V}(i) = 0, \tilde{V}(0) = 0 \qquad (3.32)$$

for $\tilde{V}(i)$ such that $\sup_{i>0} \frac{|\tilde{V}(i)|}{\eta_1^{i/n}} < \infty$.

For $i \geq k-1$, the solution is in the form

$$\tilde{V}(i) = \frac{(\tilde{\eta}+1)\eta_1^{(i+1)/n}}{(\tilde{\eta} - \eta_1^{1/n})(\eta_1^{1/n} - 1)} + Z_1, \qquad (3.33)$$

where Z_1 is a constant. Actually $\tilde{V}(i) = \frac{(\tilde{\eta}+1)\eta_1^{(i+1)/n}}{(\tilde{\eta}-\eta_1^{1/n})(\eta_1^{1/n}-1)} + Z_1 + \hat{Z}_1 \tilde{\eta}^i$ also solves the equation but is not bounded by $\eta_1^{i/n}$, meaning that $\hat{Z}_1 = 0$.

For $i \leq k$, bearing in mind that $\tilde{V}(0) = 0$, the solution is in the form

$$\tilde{V}(i) = Z_2 \tilde{\eta}^i - Z_2. \qquad (3.34)$$

In particular, we have when $i = k-1$:

$$\frac{(\tilde{\eta}+1)\eta_1^{k/n}}{(\tilde{\eta} - \eta_1^{1/n})(\eta_1^{1/n} - 1)} + Z_1 = Z_2(\tilde{\eta}^{k-1} - 1)$$

and when $i = k$:

$$\frac{(\tilde{\eta}+1)\eta_1^{(k+1)/n}}{(\tilde{\eta} - \eta_1^{1/n})(\eta_1^{1/n} - 1)} + Z_1 = Z_2(\tilde{\eta}^k - 1).$$

Solve the above two equations simultaneously and obtain:

$$\begin{aligned}
Z_1 &= \Big\{-\eta_1^{k/n}(\tilde{\eta}-\eta_1^{1/n})(\eta_1^{1/n}-1)(1-\tilde{\eta}) \\
&\quad -(1-\tilde{\eta})\eta_1^{(k+2)/n}-\eta_1^{k/n}(1-\tilde{\eta}^{k+1})(\tilde{\eta}-\eta_1^{1/n})\Big\} \\
&\quad \times\frac{1+\tilde{\eta}}{(\tilde{\eta}-\eta_1^{1/n})(\eta_1^{1/n}-1)(1-\tilde{\eta})\tilde{\eta}^{k+1}}
\end{aligned} \tag{3.35}$$

and

$$\begin{aligned}
Z_2 &= \Big\{-\eta_1^{k/n}(\tilde{\eta}-\eta_1^{1/n})(\eta_1^{1/n}-1)(1-\tilde{\eta}) \\
&\quad -(1-\tilde{\eta})\eta_1^{(k+2)/n}-\eta_1^{k/n}(1-\tilde{\eta}^{k+1})(\tilde{\eta}-\eta_1^{1/n})\Big\} \\
&\quad \times\frac{1+\tilde{\eta}}{(\tilde{\eta}-\eta_1^{1/n})(\eta_1^{1/n}-1)(1-\tilde{\eta})\tilde{\eta}^{k+1}(\tilde{\eta}^k-1)} \\
&\quad +\frac{(\tilde{\eta}+1)\eta_1^{(k+1)/n}}{(\tilde{\eta}-\eta_1^{1/n})(\eta_1^{1/n}-1)(\tilde{\eta}^k-1)}
\end{aligned}$$

Particularly, putting Z_1 given by (3.35) into the expression (3.33) and setting $i=k$ returns

$$\tilde{V}(k) = \frac{\eta_1^{k/n}(1+\tilde{\eta})(1-\tilde{\eta}^k)}{(\tilde{\eta}-\eta_1^{1/n})(1-\tilde{\eta})\tilde{\eta}^{k-1}}.$$

■

Chapter 4

Fluid approximation: EOQ and EPQ models

4.1 Introduction

As mentioned in Chapter 1, results obtained in Chapter 3 have implications on stochastic problems with a long run average expected cost. One example of such problems is the EOQ model, which we study in this chapter. In what follows, we shall accept the convention that tcu (resp. TCU) stands for total cost per unit time in fluid models (resp. stochastic models).

The classic EOQ model aiming to provide the minimizer of tcu, termed as EOQ, is primarily based on the following set of assumptions: (i) instantaneous holding cost being linear in the inventory level; (ii) constant setup cost; (iii) demand coming in a deterministic and continuous process at a constant rate; (iv) no backlogging; (v) homogeneous and nonperishable goods; (vi) continuous reviews of the inventory level; and (vii) instantaneous replenishment after ordering. Amongst the efforts of generalizing the classic EOQ models, a significantly great deal have been made on relaxing assumption (iii): for instance, Baker and Urban (1988) with a demand rate as a polynomial function of inventory level, Datta and Pal (1990); Giri et al (1996) with a demand rate discontinuous in inventory level, Urban (2005) with an initial inventory level-dependent demand rate, and Berman and Perry (2006) with demand rate of a more general inventory level-dependence, to list some. This fact is a response to: "At times, the presence of inventory has a motivational effect on the people around it. It is a common belief that large piles of goods displayed in a supermarket will lead the customer to buy more (Levin et al 1972)." Generalizations regarding other assumptions include allowing backlogging and period reviews in Urban (1995), accounting for perishable goods in Ferguson et al (2007); Giri and Chaudhuri (1998); Giri et al (1996); Weiss

(1982), and so on. A comprehensive review of the literature on EOQ models is available in Urban (2005). Note, the aforementioned works mainly focus on deterministic EOQ models.

Clearly at least to some extent, the real life is of stochastic nature, and these deterministic models can therefore be regarded as fluid approximations to the corresponding more difficult-solving stochastic EOQ models, which can arise from relaxing assumption (iii) to allow demand coming in a stochastic process. For this reason, below we shall call the deterministic models also fluid models. Suppose now the EOQ for the fluid model is obtained, naturally the next question is about how to translate the (fluid) EOQ to an order quantity for the corresponding stochastic model with a satisfactory performance, measured by TCU. As mentioned in Chapter 1, here "satisfactory" is understood in the sense of fluid scaling. Having in mind Theorem 3.1, one natural translation could be via (3.12) with a minor modification, which was indeed considered in Piunovskiy (2009a). There by modelling the underlying stochastic model as a Markovian system continuous in time, the author justified this natural translation of the (fluid) EOQ, by showing that it leads to an AFO (asymptotic fluid optimal) and AO (asymptotic optimal) (see Gajrat and Hordijk (2000); Gajrat et al (2003)) order quantity for the stochastic model[1]. Nonetheless, fairly strict conditions were assumed there, restricting the applicability of its results. By the way, the Markovian feature of the underlying system comes from assuming exponentially distributed inter-arrival time in the demand process as well as the production time, which is typical and common in the current literature about inventory systems, see He and Jewkes (2002); Köchel (1996); Weiss (1982); Xu and Chao (2009); Zheng and Zipkin (1990), for example.

Therefore, the main objective in this chapter is to justify this natural translation and provide its efficiency, but under significantly less restrictive conditions compared to those in Piunovskiy (2009a). In a nutshell, the current chapter allows the demand rate and holding cost rate to be of a rather general inventory level-dependence, so that our results are applicable to the important case of discontinuous demand rate and holding cost rate (compared to globally Lipschitz continuous rates in Piunovskiy (2009a)). Additionally, by replacing ordering from suppliers with producing gradually as the means of inventory backup, we shall also study the similar issues for EPQ models.

The rest of this chapter is organized as follows: in Section 4.2 and Section 4.3 we formulate EOQ and EPQ models and state the main results. In Section 4.4 some comments are given on the issues of possible applicabilities of the obtained results. Finally we finish this chapter with conclusions. The proofs of main statements are collected at the end in Section 4.6.

Throughout this chapter, the context should always make it clear when $[\cdot]$ means the function taking the integer part of its argument.

[1] Here we deem it more appropriate to postpone the definitions of AFO and AO order quantities to Sections 4.2 and 4.3

4.2 EOQ models

4.2.1 Description of mathematical models

Fluid model: Suppose some positive real order quantity $\varphi > 0$ is fixed, and let $\{y(t), t \geq 0\}$ be the inventory level process, whose dynamics, given the initial state $y(0) = \varphi$, is subject to $\frac{dy}{dt} = -\mu(y)$ when $0 < y(t) \leq \varphi$; and $y(t_{order} + 0) = \varphi$, where $y(t)$ reaches state zero at t_{order}. Here, $\mu(y) > 0$ is the demand rate, and the impulsive jump of $y(t)$ at t_{order} reflects the instantaneous replenishment assumption. Let t_{cycle} be the time duration between two consecutive jumps of the inventory level process, $g(y)$ the instantaneous holding cost rate, and $K > 0$ the setup cost, incurred immediately whenever an order is made. Then $tcu(\varphi) \triangleq \lim_{T\to\infty} \frac{1}{T}\left\{\int_0^T g(y(t))dt + K\left[\frac{T}{t_{cycle}}\right]\right\}$. Let us call φ^* the EOQ for the fluid model, so that $tcu(\varphi^*) = \inf_{\varphi>0} tcu(\varphi)$.

(Scaled) stochastic model: Suppose some positive integer order quantity ${}^n\Phi > 0$ is fixed. Then quite formally, the inventory level process $\{{}^nY_t, t \geq 0\}$ is modelled as a continuous time Markov chain with state space $\{0, 1, 2, \ldots, {}^n\Phi\}$ and transition rates given by

$$
{}^nq(j|0) = \begin{cases} n\kappa, & \text{if } j = {}^n\Phi; \\ -n\kappa, & \text{if } j = 0; \\ 0, & \text{otherwise}; \end{cases}
$$

in case of $i \in \{1, 2, 3, \ldots, {}^n\Phi\}$

$$
{}^nq(j|i) = \begin{cases} n\mu(\frac{i}{n}), & \text{if } j = i-1; \\ -n\mu(\frac{i}{n}), & \text{if } j = i; \\ 0, & \text{otherwise}, \end{cases}
$$

where $n\mu(\frac{i}{n})$ is the instantaneous demand rate, and $n\kappa$ is the parameter of the exponentially distributed lead time between the ordering and the corresponding replenishment. Then we have the performance functional

$$
{}^nTCU({}^n\Phi) \triangleq \lim_{T\to\infty} \frac{1}{T} E_{{}^n\Phi}\left[\int_0^\infty \left\{g(\frac{{}^nY_t}{n}) + Kn\mu(\frac{1}{n})I\{{}^nY_t = 1\}\right\} dt\right],
$$

where $E_{{}^n\Phi}$ denotes the expectation operator with initial position ${}^nY_0 = {}^n\Phi$, though indeed we can put any state as the subindex in the above definition. Let us call ${}^n\Phi^*$ the EOQ for the (scaled) stochastic model, if ${}^nTCU({}^n\Phi^*) = \inf_{{}^n\Phi=1,2,\ldots} {}^nTCU({}^n\Phi)$. Clearly, we could have formulated this optimization problem in the form of a CTMDP as in preliminary example (b). However, it appears more convenient to formulate it in the current way. Accordingly, we have divided the cost rate into two parts: the holding cost g and the setup cost K.

As in Chapter 3, let us explain the fluid scaling through the parameter $n \in \mathbb{N}$, if we put ${}^n\Phi \triangleq [n\varphi]$, where φ is the order quantity for the fluid model. Clearly, when

$n = 1$, the stochastic model is a natural corresponding version of the fluid model. As we increase n, the demand comes in smaller units, or say equivalently, inventories are measured more accurately. Take $\mu(i/n) \triangleq M$ as an example: if $n = 1$ corresponds to the unit of (kg) so that $n\mu(\frac{i}{n}) \triangleq nM = M$ (kg) comes per time unit; then in case of $n = 1000$, the unit will be (g): $n\mu(\frac{i}{n}) = 1000M$ (g) comes per time unit. Meanwhile, the cost rate is not amplified by multiplying n because it costs the same to hold either $1000M$ (g) or M (kg) of inventories.

Finally, suppose the fluid model is solved, and one obtains the (fluid) EOQ φ^*. In line with the aforementioned physical meaning of the fluid scaling, a natural translation of this (fluid) EOQ to an order quantity for the (scaled) stochastic model will be via ${}^n\Phi = [n\varphi^*]$ (see (3.12)). We say $[n\varphi]$ is AFO if $\lim_{n\to\infty} |{}^nTCU([n\varphi]) - tcu(\varphi^*)| = 0$, and AO if $\lim_{n\to\infty} |{}^nTCU([n\varphi]) - {}^nTCU({}^n\Phi^*)| = 0$. In what follows, EOQ for both fluid and stochastic models will be assumed unique, and the similar assumption applies to the EPQ models.

4.2.2 Main results

Conditions 4.1 *(a) There exist constants $d_1 \geq 0$, $k_1 > 0$ and $\delta > 0$ such that $\delta \leq \mu(y) \leq k_1$ and $|g(y)| \leq d_1$; here functions μ and g are measurable, and both defined on $[0,\infty)$.*

(b) There exist finite intervals $(0,y_1)$, (y_1,y_2), ... with $\lim_{j\to\infty} y_j = \infty$ such that on each of them, $g(y)$ and $\mu(y)$ are Lipschitz continuous, with uniformly bounded Lipschitz constants d_g and d_μ.

Note, Conditions 4.1 imply that for any fixed $\varphi > 0$, there exist L (possibly φ-dependent) finite intervals $(0,y_1)$, (y_1,y_2), ..., $(y_L, 3[\varphi+1])$ such that on each interval, $\frac{1}{\mu(y)}$ and $\frac{g(y)}{\mu(y)}$ are Lipschitz continuous with Lipschitz constants $\frac{d_\mu}{\delta^2}$ and $\frac{d_1 d_\mu + k_1 d_g}{\delta^2}$, respectively. For simplicity, we define $d_2 \triangleq \max\{\frac{d_\mu}{\delta^2}, \frac{d_1 d_\mu + k_1 d_g}{\delta^2}\}$.

Proposition 4.1 *Under Conditions 4.1, for any fixed order quantity in the fluid model $\varphi > 0$*

$$
\begin{aligned}
|{}^nTCU([n\varphi]) - tcu(\varphi)| \;\leq\; & \frac{(1+d_1)k_1^2\kappa}{\delta([n\varphi]\kappa + k_1)} \left\{ \frac{k_1 d_2 3[\varphi+1]}{\delta} + (1+\frac{k_1}{\delta})\frac{3d_1 L}{\delta} \right. \\
& \left. + \max\{\frac{g(0)}{\kappa}, \frac{1}{\kappa}\} \right\}. \qquad (4.1)
\end{aligned}
$$

In particular, Proposition 4.1 implies $\lim_{n\to\infty} |{}^nTCU([n\varphi^*]) - tcu(\varphi^*)| = 0$, i.e., $[n\varphi^*]$ is AFO.

The following lemma is from Piunovskiy (2009a), where under Conditions 4.1 and $g(y) \geq 0$, it was shown that $\varphi^* \geq \frac{\delta^2 K}{k_1 d_1}$ and ${}^n\Phi^*$ satisfies $\frac{{}^n\Phi^*}{n} \geq \frac{\delta^2 K}{k_1 d_1} - \frac{\delta}{n\kappa}$.

Lemma 4.1 *Under Conditions 4.1, if $g(y) \geq 0$, then $\liminf_{n\to\infty} \frac{{}^n\Phi^*}{n} > 0$.*

Corollary 4.1 *Under Conditions 4.1, if* $\liminf_{n\to\infty} \frac{^n\Phi^*}{n} > 0$ *(see Lemma 4.1), then we have* $\lim_{n\to\infty} |\frac{^n\Phi^*}{n} - \frac{[n\varphi^*]}{n}| = 0$, *and* $\lim_{n\to\infty} |^nTCU([n\varphi^*]) - {}^nTCU(^n\Phi^*)| = 0$, *i.e.,* $[n\varphi^*]$ *is AO.*

Regarding Proposition 4.1, as $\kappa \to \infty$, which corresponds to the case of no lead time, the RHS of (4.1) goes to $\frac{(1+d_1)k_1^2}{\delta[n\varphi]} \left\{ \frac{k_1 d_2 3[\varphi+1]}{\delta} + (1+\frac{k_1}{\delta})\frac{3d_1L}{\delta} \right\}$, which blows up to ∞ as φ goes to 0, while all the other parameters are fixed. This raises our desirability for a φ-independent estimate, which can be done with some additional conditions, as given by the following corollary.

Corollary 4.2 *Under Conditions 4.1, suppose in addition L is* φ*-independent and* $g(y) \geq 0$ *on* $[0,\infty)$*. Then the following statements hold:*
(a) For any fixed $n \geq N$ *where N is any positive integer number satisfying* $N > \frac{(2+\frac{\delta}{\kappa})k_1 d_1}{\delta^2 K}$,
and for any $\varphi \geq \frac{\delta^2 K}{k_1 d_1} - \frac{\delta}{n\kappa}$,

$$\begin{aligned}
|^nTCU([n\varphi]) - tcu(\varphi)| \leq\ & \frac{3(1+d_1)k_1^3 d_2}{\delta^2} \frac{1}{n\left\{1 - \frac{\kappa k_1 d_1}{n\kappa\delta^2 K - \delta k_1 d_1}\right\}} \\
& + \frac{(1+d_1)k_1^2}{n\delta\left\{\frac{\delta^2 K}{k_1 d_1} - \frac{\delta}{n\kappa} - \frac{1}{n}\right\}} \\
& \times \left\{\frac{3k_1 d_2}{\delta} + (1+\frac{k_1}{\delta})\frac{3d_1 L}{\delta} + \max\{\frac{g(0)}{\kappa}, \frac{1}{\kappa}\}\right\} \triangleq E(n).
\end{aligned}$$

Here, $E(n)$ *is* φ*-independent and* $E(n) = O(\frac{1}{n})$.
(b) For any $n \geq N$,

$$|^nTCU(^n\Phi^*) - {}^nTCU([n\varphi^*])| \leq 2E(n) = O(\frac{1}{n}).$$

Here we recall that $^n\Phi^*$ *and* φ^* *are the EOQ for the (scaled) stochastic model and fluid model, respectively.*

Corollary 4.2 refines Corollary 4.1 for certain cases by providing an estimate of the rate of convergence.

Fixing some order quantity for the fluid model φ and scaling parameter n, and let $^n\Pi(i)$, $i = 0, 1, \ldots [n\varphi]$ be the stationary distribution of the inventory process $\{^nY_t, t \geq 0\}$, and $\pi(y)$ the invariant density in the fluid model of the underlying dynamics of $y(t)$. Here we put $\pi(0) = 0$ for convenience. Then the following proposition shows that the fluid model can also be used to provide approximations to stationary distributions of the inventory level process in the (scaled) stochastic model.

Proposition 4.2 *Suppose Conditions 4.1 are satisfied, and some order quantity for the* *fluid model $\varphi > 0$ is fixed. Then we have*

$$|^n\Pi(0) - \pi(0)| \le \frac{k_1}{k_1 + \kappa[n\varphi]};$$

and for $i = 1, 2, \ldots, [n\varphi]$

$$|^n\Pi(i) - \int_{\frac{i-1}{n}}^{\frac{i}{n}} \pi(y)dy| \quad \le \quad \frac{\frac{2\varphi k_1^2\kappa}{\delta^2} + \frac{k_1^2\kappa}{n\delta}(\frac{3k_1 d_2(\varphi+1)}{\delta} + (1+\frac{k_1}{\delta})\frac{3d_1 L}{\delta} + \frac{1}{\kappa})}{([n\varphi]+1)\varphi}.$$

4.3 EPQ models

In this section, results similar to those derived in the previous section are produced for EPQ models. So we shall omit any repetitive comments.

4.3.1 Description of mathematical models

Fluid model: Suppose we fix some real inventory backup level $\varphi > 0$, meaning that once it is switched on, the production is always on until the inventory level reaches φ. Let $\{y(t), t \ge 0\}$ represent the inventory level process in the fluid model, with state space $[0, \varphi]$, and instantaneous demand and production rate $\mu(y) > 0$ and $\lambda(y) > 0$, respectively. Then given the initial state $y(0) = \varphi$, the inventory level process is subject to the dynamics

$$\frac{dy}{dt} = \begin{cases} -\mu(y) & \text{(production-off phase)}; \\ \lambda(y) - \mu(y) & \text{(production-on phase)}, \end{cases}$$

where the production-off phase and production-on phase, superseding each other, are triggered by $y(t) = \varphi$ and $y(t) = 0$, respectively. In words, without any delay, once the inventory level reaches zero, production is switched on till it reaches the inventory backup level φ. Let $g(y)$ be the holding cost rate, $K > 0$ the setup cost incurred with switching on the production, and t_{cycle} the time duration between two consecutive production switching-offs. So $tcu(\varphi) \triangleq \frac{1}{T}\lim_{T\to\infty}\left\{\int_0^T g(y(t))dt + K\left[\frac{T}{t_{cycle}}\right]\right\}$. Let us denote φ^* the EIBL[2] (Economic Inventory Backup Level) for the fluid model, so that $tcu(\varphi^*) = \inf_{\varphi>0} tcu(\varphi)$.

(Scaled) stochastic model: Suppose we fix some inventory backup level positive integer $^n\Phi$, meaning that the production is always on until the inventory reaches the level $^n\Phi$. Let $\{^nY_t, t \ge 0\}$ represent the inventory level process. We shall model it as a continuous time Markov chain with the state space

$$\{(^n\Phi, off), (^n\Phi - 1, off), \ldots, (0, off), (0, on), \ldots, (^n\Phi - 1, on)\},$$

[2]Note, here we are interested in EIBL, rather than the economic production lot size.

where (i,off) indicates that the inventory level is i and the production is off; the denotation of (i,on) can be understood in the same way. Its transition rates are given by

$$
\begin{aligned}
&{}^{n}q((0,on)|(0,off)) = n\kappa,\ {}^{n}q((0,off)|(0,off)) = -n\kappa;\\
&{}^{n}q((1,on)|(0,on)) = n\lambda(0),\ {}^{n}q((0,on)|(0,on)) = -n\lambda(0);\\
&{}^{n}q(({}^{n}\Phi,off)|({}^{n}\Phi-1,on)) = n\lambda(\frac{{}^{n}\Phi-1}{n}),\\
&{}^{n}q(({}^{n}\Phi-2,on)|({}^{n}\Phi-1,on)) = n\mu(\frac{{}^{n}\Phi-1}{n}),\\
&{}^{n}q(({}^{n}\Phi-1,on)|({}^{n}\Phi-1,on)) = -n\lambda(\frac{{}^{n}\Phi-1}{n}) - n\mu(\frac{{}^{n}\Phi-1}{n});
\end{aligned}
$$

$\forall i = 1,\ldots,{}^{n}\Phi$:

$$
{}^{n}q((j,off)|(i,off)) = \begin{cases} n\mu(\frac{i}{n}), & \text{if } j = i-1;\\ -n\mu(\frac{i}{n}), & \text{if } j = i; \end{cases}
$$

and finally $\forall i = 1,\ldots,{}^{n}\Phi-2$:

$$
{}^{n}q((j,on)|(i,on)) = \begin{cases} n\mu(\frac{i}{n}), & \text{if } j = i-1;\\ n\lambda(\frac{i}{n}), & \text{if } j = i+1;\\ -n\mu(\frac{i}{n}) - n\lambda(\frac{i}{n}), & \text{if } j = i, \end{cases}
$$

where $n\lambda(\frac{i}{n})$ and $n\mu(\frac{i}{n})$ stand for the instantaneous production and demand rates, $n\kappa$ is the parameter of the exponentially distributed lead time between the switching and the actual production-on, and we have ignored all the cases when the transition rates take zero. So we have ${}^{n}TCU({}^{n}\Phi) \triangleq \lim_{T\to\infty}\frac{1}{T}E_{{}^{n}\Phi}[\int_0^T \left\{g(\frac{{}^{n}Y_t}{n}) + Kn\mu(\frac{1}{n})I\{{}^{n}Y_t = 1\}\right\}dt]$, with the holding cost $g(\frac{i}{n})$ and setup cost $K > 0$. Let us denote ${}^{n}\Phi^*$ the EIBL for the (scaled) stochastic model, so that ${}^{n}TCU({}^{n}\Phi^*) = \inf_{{}^{n}\Phi=1,2,\ldots}{}^{n}TCU({}^{n}\Phi)$.

The concept of AFO and AO inventory backup level $[n\varphi]$ can be understood in the same manner as introduced at the end of Subsection 4.2.1.

4.3.2 Main results

Conditions 4.2 *(a) There exist some constants $d_1 \geq 0$, $k_1 > 0$, $\delta > 0$, $\delta_{\mu\lambda} > 0$, $\tilde{\eta} > 1$ and measurable functions $\mu(y)$, $\lambda(y)$, and $g(y)$ defined on $[0,\infty)$ such that $\delta \leq \mu(y) \leq \mu(y) + \delta_{\mu\lambda} \leq \lambda(y) \leq k_1$, $\lambda(y) + \mu(y) \leq k_1$, $|g(y)| \leq d_1$ and $\inf_{y>0}\frac{\lambda(y)}{\mu(y)} = \tilde{\eta}$.*
(b) There exist finite intervals $(0,y_1)$, (y_1,y_2), ... with $\lim_{j\to\infty} y_j = \infty$ such that on each of them $\lambda(y)$, $\mu(y)$ and $g(y)$ are Lipschitz continuous with uniformly bounded Lipschitz constants d_λ, d_μ and d_g, respectively.

Note, Conditions 4.2 imply that for any fixed $\varphi > 0$, there exists an integer L (possibly φ-dependent) and $L+1$ intervals $(0,y_1)$, (y_1,y_2), ..., $(y_L,3[\varphi+1])$ such that on each interval $\frac{1}{\mu(y)}$, $\frac{g(y)}{\mu(y)}$, and $\frac{1}{\lambda(y)-\mu(y)}$ are Lipschitz continuous with Lipschitz constant $\frac{d_\mu}{\delta^2}$, $\frac{k_1 d_g + d_1 d_\mu}{\delta^2}$ and $\frac{d_\lambda + d_\mu}{\delta^2_{\mu\lambda}}$, respectively; and at the same time on each of the intervals, functions (with respect to y) $\frac{g(\varphi-y)}{\lambda(\varphi-y)-\mu(\varphi-y)}$ and $\frac{1}{\lambda(\varphi-y)-\mu(\varphi-y)}$ are Lipschitz continuous with Lipschitz constants $\frac{k_1 d_g + d_1 d_\mu}{\delta^2}$ and $\frac{d_\lambda + d_\mu}{\delta^2_{\mu\lambda}}$, respectively. Let us now denote the common Lipschitz constant as $d_2 \triangleq \max\{\frac{d_\mu}{\delta^2}, \frac{k_1 d_g + d_1 d_\mu}{\delta^2}, \frac{d_\lambda + d_\mu}{\delta^2_{\mu\lambda}}\}$.

Proposition 4.3 *Under Conditions 4.2, for any fixed (fluid) inventory backup level* $\varphi > 0$,

$$
\begin{aligned}
& |^n TCU([n\varphi]) - tcu(\varphi)| \\
\leq \ & \frac{2\bar{\delta}\kappa k_1(1+d_1)}{\underline{\delta}} \left\{ \frac{B_1 + B_2 n \tilde{\eta}^{-2n[\varphi+1]} + \max\{\frac{g(0)}{2\kappa}, \frac{1}{2\kappa}\}}{2[n\varphi]\kappa + k_1} \right\},
\end{aligned}
$$

where we put $\bar{\delta} \triangleq \max\{\delta, \ \delta_{\mu\kappa}\}$ *and* $\underline{\delta} \triangleq \min\{\delta, \ \delta_{\mu\kappa}\}$, *and n-independent* B_1, B_2 *are given in (4.3), (4.4) in Section 4.6.*

As in the case of EOQ models, we observe from Proposition 4.3 that $[n\varphi^*]$ is AFO.

Corollary 4.3 *Under Conditions 4.2, the following statements hold:*
(a) If $g(y) \geq 0$, *then* $\liminf_{n\to\infty} \frac{^n\Phi^*}{n} > 0$.
(b) If $\liminf_{n\to\infty} \frac{^n\Phi^*}{n} > 0$, *then*

$$\lim_{n\to\infty} |\frac{^n\Phi^*}{n} - \frac{[n\varphi^*]}{n}| = 0$$

and

$$\lim_{n\to\infty} |^n TCU([n\varphi^*]) - {^n TCU}(^n\Phi^*)| = 0,$$

i.e., $[n\varphi^*]$ *is AO.*
(c) If L is φ*-independent and* $g(y) \geq 0$, *then the following two sub-statements hold:*
(c1) For $\varphi \geq \frac{K\delta^2}{4k_1 d_1}$ *and big enough n so that* $n > \frac{4k_1 d_1}{K\underline{\delta}^2}$ *and* $n > \frac{k_1}{2\kappa} - 1$,

$$
\begin{aligned}
& |^n TCU([n\varphi]) - tcu(\varphi)| \\
\leq \ & \frac{2\bar{\delta}\kappa k_1(1+d_1)}{\underline{\delta}} \left\{ \frac{\hat{B}_1 + \hat{B}_2 n \tilde{\eta}^{-2n(\frac{K\delta^2}{4k_1 d_1}+1)} + \max\{\frac{g(0)}{2\kappa}, \frac{1}{2\kappa}\}}{2(n(\frac{K\delta^2}{4k_1 d_1}) - 1)\kappa + k_1} \right\} \\
\triangleq \ & F(n) = O(\frac{1}{n}).
\end{aligned}
$$

Here $F(n)$ is φ-independent, and

$$\hat{B}_1 = \frac{k_1 d_2(\tilde{\eta}+1)3(\frac{K\delta^2}{4k_1d_1}+1)}{\delta(\tilde{\eta}-1)} + \left(1+\frac{k_1\tilde{\eta}}{\delta(\tilde{\eta}-1)}\right)\frac{3d_1L(\tilde{\eta}+1)}{\delta(\tilde{\eta}-1)},$$

$$\hat{B}_2 = \left(1+\frac{6k_1(\frac{K\delta^2}{4k_1d_1}+1)\tilde{\eta}}{\delta(\tilde{\eta}-1)}\right)\frac{d_1\tilde{\eta}^2(\tilde{\eta}+1)}{\delta(\tilde{\eta}-1)^2}.$$

(c2) For sufficiently big n,

$$|{}^nTCU({}^n\Phi^*) - {}^nTCU([n\varphi^*])| \leq 2F(n) = O(\frac{1}{n}).$$

As a case study, let us consider the following classic setting for the (scaled) stochastic EPQ model: assume constant demand and production rates $\mu(y) \triangleq D > 0$ and $\lambda(y) \triangleq R > 0$, linear holding cost $g(y) = hy$ with a constant $h > 0$, constant setup cost $K > 0$, and finally no lead time between "switching" and "actual production-on", corresponding to if we take $\kappa \to \infty$. Therefore, if we consider the underlying continuous time Markov chain $\{{}^nY_t, t \geq 0\}$, state $(0, off)$ will be excluded. The transition rates are modified accordingly. Clearly, Conditions 4.2 are satisfied with this classic setting. The following lemma gives the explicit formula for ${}^nTCU(Z)$, where for simplicity we have put Z instead of ${}^n\Phi$ for the inventory backup level.

Proposition 4.4

$$\begin{aligned}{}^nTCU(Z) \;=\; & \{RhZ^2(D-R)^2 + 2Kn^2D(R-D)^3 + ZhR(3D-R)(D-R) \\ & +2RhD^2(1-(\frac{D}{R})^Z)\} \\ & \times \frac{1}{2n(R-D)\{ZR^2 - D^2 + D^2(\frac{D}{R})^Z - ZRD\}}.\end{aligned}$$

Clearly, if we put $Z = [n\varphi]$ in the expression for ${}^nTCU(Z)$, with φ the inventory backup level for the fluid model, then one can easily see that $\lim_{n\to\infty} {}^nTCU([n\varphi]) = \frac{KD(R-D)}{\varphi R} + \frac{h\varphi}{2} = tcu(\varphi)$, which agrees with Proposition 4.3. Secondly, if n increases, then by inspecting the numerator (especially the first two terms) and the denominator, we see that if Z does not increase as fast as n, ${}^nTCU(Z)$ will blow up to ∞. Note, it can be easily checked that the expression $ZR^2 - D^2 + D^2(\frac{D}{R})^Z - ZRD > 0$. Therefore, the condition of $\liminf_{n\to\infty} \frac{{}^n\Phi^*}{n} > 0$ in Corollary 4.3 is satisfied.

Finally, Let us finish this section with the following remark. Although the above studied EPQ model makes sense in its own right, it is also somehow related to another interesting problem in telecommunications, the "buffer sizing" problem. For illustrative means, let us take the following naive fluid model as an example. Consider a connection with a source sending data packets to a router, which further switches the data to the sink at rate $\mu(y)$, where y is the instantaneous number of packets in the buffer. Suppose at initial time, the buffer of the router is empty. The source sends information packets to the router at rate $\lambda(y) > \mu(y)$, till the point when the state of the buffer reaches the buffer size. After that point, the sender pauses until the next moment when the buffer is empty. The problem is to size the buffer so as to minimize the average holding cost. Clearly, this is just another interpretation of the EPQ model. Surely the real traffic is of stochastic nature. However, often a fluid model is taken for analysis, see for instance, Avrachenkov et al (2005, 2010), where the authors studied the buffer sizing problem for the real Internet traffic governed by TCP implying a time- and state-dependent sending rate of data. For the network traffic governed by STCP, we shall study this buffer sizing problem in Chapter 6.

4.4 Comments and remarks

In this section, let us comment on the main results, Propositions 4.1 and 4.3, and their applicability issues. We shall mainly focus on EOQ models, as absolutely similar comments can be made on EPQ models in the same manner.

Although we assume the ordering point to be always zero, our results are still applicable when it is set to some fixed positive level, because Lemma 4.2 holds if we put another absorbing state instead of zero. In particular, if one allows the state taking negative values, by putting some negative state absorbing, our results also impound the case of backlogging. This flexibility regarding the ordering point together with the fact that g is unrestricted in signs enriches the applications of our results, in that although we require μ to be separated from zero, when profit rather than solely operational cost is counted, the ordering point is most typically positive, meaning in cases of $\mu(y) = \alpha y^{\beta}$, $\alpha > 0$, $0 < \beta < 1$ as in Baker and Urban (1988) and $\mu(y) = \alpha y^{-\beta}$, $\alpha > 0$, $\beta \geq 1$ as in Berman and Perry (2006), μ will be essentially separated from zero.

Indeed, the state-dependence given in Conditions 4.1 and 4.2 is fairly general. In particular, that functions $\mu(y)$ and $g(y)$ being bounded is not too restrictive, because once some EOQ for the fluid model $\varphi > 0$ is fixed, to validate Propositions 4.1 and 4.3, they are only required to be bounded on bounded intervals. For example, once the ordering point is positive, the state space is essentially a closed interval, and the aforementioned two forms of demand rate in Baker and Urban (1988); Berman and Perry (2006) will satisfied all the conditions required here. Note, in addition to the demand rate, some authors such as Giri and Chaudhuri (1998); Giri et al (1996) also included a state-dependent deteriorating rate, to indicate that the underlying goods are perishable. Our results are also applicable to such cases: one only needs explain μ as the reduction rate of the inventory level.

Finally, Propositions 4.1 and 4.3 are significant extensions of Piunovskiy (2009a), where the author focused on EOQ models only and required (global) Lipschitz continuity of μ and g. However, from the practical point of view, the case of discontinuous functions is interesting and important, as demonstrated by Berman and Perry (2006); Datta and Pal (1990); Giri et al (1996); Urban (1992, 1995), where Berman and Perry (2006) considered a piecewise constant function μ and the others considered μ taking either a constant value or according to $\mu(y) = \alpha y^{\beta}$, $\alpha > 0$, $0 < \beta < 1$. The results in Piunovskiy (2009a) were derived based on the closed-form of the solution to a Poisson equation satisfied by ${}^{n}TCU$, which is tremendously difficult to get explicitly in the case of (stochastic) EPQ models; into which nevertheless the current work provides insight.

4.5 Conclusion

To sum up, in this chapter we formally justified a general class of inventory level-dependent deterministic EOQ and EPQ models, regarded as the fluid approximations to their stochastic versions in that a natural translation of the fluid EOQ (EIBL) was shown to provide an order quantity (inventory backup level) for the stochastic model achieving some optimality asymptotically. The efficiency of the translation mechanism was obtained, too. The class of inventory models are quite broad so that to various extent, the obtained results are directly applicable to existing works such as Baker and Urban (1988); Berman and Perry (2006); Datta and Pal (1990); Giri and Chaudhuri (1998); Giri et al (1996); Urban (1992, 2005). This chapter can be viewed as a refinement of Piunovskiy (2009a).

4.6 Proof of main statements

To aid our proof, firstly, let us consider the following one-dimensional Birth-and-Death process $\{{}^{n}Z_t,\ t \geq 0\}$ with state space $\{0, 1, \dots\}$ and birth and death rates $n\alpha(\frac{i}{n})$ and $n\beta(\frac{i}{n})$, respectively, where nonnegative measurable functions α and β are defined on $[0, \infty)$ and i indicates the current state of the process. In addition, $\alpha(0) = \beta(0) = 0$, meaning that state zero is absorbing. Let E_i denote the expectation of any underlying functional of the process, with the initial state ${}^{n}Z_0 = i$. Let a real measurable function γ defined on $[0, \infty)$ be fixed with $\gamma(0) = 0$. Now we are in the position to state the following conditions:

Conditions 4.3 *(a) There exist constant* $\tilde{\eta} > 1$, $\delta > 0$, $d_1 > 0$ *and* $k_1 < \infty$ *such that* $\inf_{z>0} \frac{\beta(z)}{\alpha(z)} > \tilde{\eta}$, $\beta(z) \geq \delta$, $\alpha(z) + \beta(z) \leq k_1$, $|\gamma(z)| \leq d_1$. *Here, if* $\alpha(z) \equiv 0$, *then* $\tilde{\eta}$ *can be arbitrary.*
(b) There exist finite intervals $(z_0 \triangleq 0, z_1)$, $(z_1, z_2), \dots$ *with* $\lim_{j \to \infty} z_j = \infty$ *such that on each of them,* $\frac{\gamma(z)}{\beta(z) - \alpha(z)}$ *is a Lipschitz continuous function with a uniformly bounded Lipschitz constant* d_2.

Note, Conditions 4.3 (b) implies that for any fixed $\varphi > 0$ there exists an integer L (possibly φ-dependent) and $L+1$ finite intervals $(0, z_1)$, $(z_1, z_2), \ldots, (z_L, 3[\varphi+1])$ such that on each interval, function $\frac{\gamma}{\beta-\alpha}$ is Lipschitz continuous with uniformly bounded Lipschitz constant d_2.

The following lemma is a slightly stronger version of (Piunovskiy 2009b, Thm.2), and will play an important role in our proof.

Lemma 4.2 *Suppose Conditions 4.3 are satisfied. Then, for each $\varphi > 0$*

$$\sup_{0 \le i \le n[\varphi+1]} \left| E_i\left[\int_0^\infty \gamma(\frac{^nZ_s}{n})ds\right] - \int_0^\infty \gamma(z(s))ds \right| \le \frac{B_1}{n} + B_2 \tilde{\eta}^{-2n[\varphi+1]}, \tag{4.2}$$

where regarding the second integral the underlying dynamics is given by $\frac{dz}{ds} = \alpha(z) - \beta(z)$, $z(0) = \frac{i}{n}$; and B_1 and B_2 are given as follows:

$$B_1 = \frac{k_1 d_2(\tilde{\eta}+1)3[\varphi+1]}{\delta(\tilde{\eta}-1)} + \left(1 + \frac{k_1\tilde{\eta}}{\delta(\tilde{\eta}-1)}\right)\frac{3d_1L(\tilde{\eta}+1)}{\delta(\tilde{\eta}-1)}, \tag{4.3}$$

$$B_2 = \left(1 + \frac{6k_1[\varphi+1]\tilde{\eta}}{\delta(\tilde{\eta}-1)}\right)\frac{d_1\tilde{\eta}^2(\tilde{\eta}+1)}{\delta(\tilde{\eta}-1)^2}. \tag{4.4}$$

Proof. It can be easily checked in the proof of (Piunovskiy 2009a, Thm.2) that our Conditions 4.3, weaker than the original conditions imposed, are sufficient for the statement. See also Chapter 3. ■

4.6.1 Proof of Proposition 4.1, Corollary 4.1, Corollary 4.2 and Proposition 4.2

Let us call, for both the fluid model and (scaled) stochastic model, the time duration between two consecutive replenishments a cycle, and denote them t_{cycle} and $^nT_{cycle}$, respectively. Here, for simplicity, we do not explicitly indicate the φ-dependence ($^n\Phi$-dependence) of t_{cycle} ($^nT_{cycle}$). Clearly, $\{^nY_t, t \ge 0\}$ is a regenerative process (Ross 2002, p.425), in that it repeats probabilistically itself from one cycle to the next. It then follows from (Hu and Liu 1999, Th.1.1, p.131) (see also (Ross 2002, Prop.7.3)) that as far as the long-run average $^nTCU([n\varphi])$ is concerned, it suffices to consider the inventory level process and the cost incurred with it over one cycle. For simplicity, we shall always consider the cycle starting at time 0 with the initial position $^nY_0 = [n\varphi]$. Let us denote ^nTC and tc the total cost incurred over the cycle in the stochastic and fluid model, respectively. Then the following lemma indicates that the differences between $E_{[n\varphi]}\left[^nT_{cycle}\right]$ and t_{cycle} and between $E_{[n\varphi]}[^nTC]$ and tc cannot be too big.

Lemma 4.3 *Under Conditions 4.1, the following two inequalities hold with nonnegative B_1 and B_2 given in (4.3) and (4.4):*

(a) $\left|E_{[n\varphi]}[{}^nT_{cycle}] - t_{cycle}\right| \leq \frac{B_1}{n} + B_2\tilde{\eta}^{-n[\varphi+1]} + \frac{1}{n\kappa}$;
(b) $\left|E_{[n\varphi]}[{}^nTC] - tc\right| \leq \frac{B_1}{n} + B_2\tilde{\eta}^{-n[\varphi+1]} + \frac{g(0)}{n\kappa}$.

Proof. (a) Let us denote ${}^nT_{absorbing}$ the time duration from the starting point $t = 0$ up to the point when ${}^nY_t$ firstly reaches state 0. Obviously, we have $E_{[n\varphi]}[{}^nT_{absorbing}] = E_{[n\varphi]}\left[\int_0^\infty I\{{}^nY_t > 0\}dt\right]$. Then $E_{[n\varphi]}[{}^nT_{cycle}] = E_{[n\varphi]}[{}^nT_{absorbing}] + \frac{1}{n\kappa}$, where the second term on the right hand side is the expected lead time. Now, notice that Conditions 4.1 is a specific version of Conditions 4.3, where $\tilde{\eta} > 1$ is arbitrary, if we take functions $\alpha(y) \equiv 0$, $\beta(y) = \mu(y)$ (here it does no matter to put $\mu(0) = 0$) and $I\{y > 0\} = g(y) = \gamma(y)$; and the inventory level process from $t = 0$ up to ${}^nT_{absorbing}$ is a pure death process. Therefore, one can refer to Lemma 4.2 for $\left|E_{[n\varphi]}[{}^nT_{cycle}] - t_{cycle}\right| \leq |E_{[n\varphi]}[{}^nT_{absorbing}] - t_{cycle}| + \frac{1}{n\kappa} \leq \frac{B_1}{n} + B_2\tilde{\eta}^{-n[\varphi+1]} + \frac{1}{n\kappa}$.

(b) Let us denote ${}^nTC_{absorbing}$ the cost incurred during the interval $[0, {}^nT_{absorbing}]$, so that $E_{[n\varphi]}[{}^nTC] = E_{[n\varphi]}[{}^nTC_{absorbing}] + \frac{g(0)}{n\kappa}$, where the second term on the right hand side corresponds to the cost incurred over the lead time. In the same way as in part (a), comparing $E_{[n\varphi]}[{}^nTC_{absorbing}]$ with tc first, and then adding $\frac{g(0)}{n\kappa}$ results in the statement. Remember, the setup cost cancels out. ■

Proof of Proposition 4.1. Under Conditions 4.1, we have

$$\begin{aligned}
& |{}^nTCU([n\varphi]) - tcu(\varphi)| \\
= & \left|\frac{E_{[n\varphi]}[{}^nTC]}{E_{[n\varphi]}[{}^nT_{cycle}]} - \frac{tc}{t_{cycle}}\right| \\
= & \left|\frac{E_{[n\varphi]}[{}^nTC]t_{cycle} - tcE_{[n\varphi]}[{}^nT_{cycle}]}{E_{[n\varphi]}[{}^nT_{cycle}]t_{cycle}}\right| \\
= & \left|\frac{E_{[n\varphi]}[{}^nTC]t_{cycle} - tc\, t_{cycle} + tc\, t_{cycle} - tcE_{[n\varphi]}[{}^nT_{cycle}]}{E_{[n\varphi]}[{}^nT_{cycle}]t_{cycle}}\right| \\
= & \left|\frac{t_{cycle}\{E_{[n\varphi]}[{}^nTC] - tc\} + tc\{t_{cycle} - E_{[n\varphi]}[{}^nT_{cycle}]\}}{E_{[n\varphi]}[{}^nT_{cycle}]t_{cycle}}\right| \\
\leq & \frac{t_{cycle}|E_{[n\varphi]}[{}^nTC] - tc| + tc|t_{cycle} - E_{[n\varphi]}[{}^nT_{cycle}]|}{E_{[n\varphi]}[{}^nT_{cycle}]t_{cycle}} \qquad (4.5) \\
\leq & \frac{(\frac{\varphi}{\delta} + d_1\frac{\varphi}{\delta})}{\frac{\varphi}{k_1}(\frac{[n\varphi]}{nk_1} + \frac{1}{n\kappa})}\left\{\frac{B_1}{n} + B_2\tilde{\eta}^{-2n[\varphi+1]} + \max\{\frac{1}{n\kappa}, \frac{g(0)}{n\kappa}\}\right\} \\
= & \frac{(1+d_1)nk_1^2\kappa}{\delta([n\varphi]\kappa + k_1)}\left\{\frac{B_1}{n} + B_2\tilde{\eta}^{-2n[\varphi+1]} + \max\{\frac{1}{n\kappa}, \frac{g(0)}{n\kappa}\}\right\},
\end{aligned}$$

where the last inequality follows from the facts that $\frac{\varphi}{k_1} \leq t_{cycle} \leq \frac{\varphi}{\delta}$, $tc \leq \frac{d_1\varphi}{\delta}$, $E_{[n\varphi]}[{}^nT_{cycle}] \geq \frac{[n\varphi]}{nk_1} + \frac{1}{n\kappa}$ and finally Lemma 4.3.

Now let us easily observe that $\frac{\tilde{\eta}+1}{\tilde{\eta}-1}$ and $\frac{\tilde{\eta}}{\tilde{\eta}-1}$ both decrease with $\tilde{\eta} \in (1, \infty)$. Then it follows that $\frac{B_1}{n}$, $B_2\tilde{\eta}^{-2n[\varphi+1]} = \frac{B_2}{\tilde{\eta}}\tilde{\eta}^{1-2n[\varphi+1]}$, and thus the above derived expression all decrease with $\tilde{\eta}$. Here we recall that $\tilde{\eta}$ can be an arbitrary number on the interval

$(1,\infty)$, see Conditions 4.3. This implies that

$$
\begin{aligned}
& |^{n}TCU([n\varphi]) - tcu(\varphi)| \\
\leq & \lim_{\tilde{\eta}\to\infty} \frac{(1+d_1)k_1^2 n\kappa}{\delta([n\varphi]\kappa+k_1)} \left\{ \frac{B_1}{n} + B_2\tilde{\eta}^{-2n[\varphi+1]} + \max\{\frac{1}{n\kappa}, \frac{g(0)}{n\kappa}\} \right\} \\
= & \frac{(1+d_1)k_1^2\kappa}{\delta([n\varphi]\kappa+k_1)} \left\{ \frac{k_1 d_2 3[\varphi+1]}{\delta} + (1+\frac{k_1}{\delta})\frac{3d_1 L}{\delta} + \max\{\frac{g(0)}{\kappa}, \frac{1}{\kappa}\} \right\}.
\end{aligned}
$$

■

Proof of Corollary 4.1. For any fixed n, let us denote $^{n}\Phi^* = n\hat{\varphi}(n)$. We shall do the proof in two parts.

(a) Firstly we consider the case of a convergent sequence $\hat{\varphi}(n)$. Suppose now that as $n \to \infty$, $\hat{\varphi}(n)$ does not go to φ^* but $\lim_{n\to\infty}\hat{\varphi}(n) = \tilde{\varphi} > 0$; here we allow $\tilde{\varphi}$ being from the extended real line. In particular, for big enough n, $\hat{\varphi}(n)$ is separated from zero. According to the proof of Proposition 4.1, we have $^{n}TCU(n\hat{\varphi}(n)) \to tcu(\hat{\varphi}(n))$. But we also have $tcu(\hat{\varphi}(n)) \to tcu(\tilde{\varphi})$, since $tcu(\varphi) = \frac{\int_0^{\varphi} \frac{g(y)}{\mu(y)}dy+K}{\int_0^{\varphi}\frac{dy}{\mu(y)}}$ is continuous in φ. This gives $^{n}TCU(n\hat{\varphi}(n)) \to tcu(\tilde{\varphi}) > tcu(\varphi^*)$. However, it follows from Proposition 4.1 that $^{n}TCU([n\varphi^*]) \to tcu(\varphi^*)$. This indicates that at least for big enough n, $^{n}TCU([n\varphi^*]) < {}^{n}TCU(n\hat{\varphi}(n)) = {}^{n}TCU(^{n}\Phi^*)$, which is a desired contradiction. Hence $\lim_{n\to\infty}\hat{\varphi}(n) = \varphi^*$, and consequently, $\lim_{n\to\infty} |^{n}TCU([n\varphi^*]) - {}^{n}TCU(^{n}\Phi^*)| = 0$, as required.

(b) Now consider the case of a divergent sequence $\hat{\varphi}(n)$. One only needs consider the following two situations: either it has a bounded subsequence, which, by Bolzano-Weierstrass' theorem, further has a convergent subsequence; or it does not have a bounded subsequence, which means that it has a subsequence blowing up to ∞. However, by taking the corresponding subsequence, we find that both situations have been essentially covered in part (a). Part (b) is thus proved. ■

Proof of Corollary 4.2. (a) For any fixed $n \geq N$ such that $N(\frac{\delta^2 K}{k_1 d_1} - \frac{\delta}{N\kappa}) > 2 \Leftrightarrow N > \frac{(2+\frac{\delta}{\kappa})k_1 d_1}{\delta^2 K}$, for any $\varphi \geq \frac{\delta^2 K}{k_1 d_1} - \frac{\delta}{n\kappa}$, we have

$$
\begin{aligned}
& \text{RHS of (4.1)} \\
\leq & \frac{(1+d_1)k_1^2}{\delta[n\varphi]} \left\{ \frac{k_1 d_2(3\varphi+3)}{\delta} + (1+\frac{k_1}{\delta})\frac{3d_1 L}{\delta} + \max\{\frac{g(0)}{\kappa}, \frac{1}{\kappa}\} \right\} \\
= & \frac{(1+d_1)k_1^3 3d_2\varphi}{\delta^2[n\varphi]} + \frac{(1+d_1)k_1^2}{\delta[n\varphi]} \left\{ \frac{3k_1 d_2}{\delta} + (1+\frac{k_1}{\delta})\frac{3d_1 L}{\delta} + \max\{\frac{g(0)}{\kappa}, \frac{1}{\kappa}\} \right\}
\end{aligned}
$$

$$\leq \frac{3(1+d_1)k_1^3 d_2}{\delta^2}\frac{\varphi}{n\varphi-1}+\frac{(1+d_1)k_1^2}{\delta(n\varphi-1)}\left\{\frac{3k_1d_2}{\delta}+(1+\frac{k_1}{\delta})\frac{3d_1L}{\delta}\right.$$
$$\left.+\max\{\frac{g(0)}{\kappa},\frac{1}{\kappa}\}\right\}$$

(Recall, here $n\varphi-1>0$.)

$$\leq \frac{3(1+d_1)k_1^3 d_2}{\delta^2}\frac{1}{n\left\{1-\frac{\kappa k_1 d_1}{n\kappa\delta^2K-\delta k_1 d_1}\right\}}+\frac{(1+d_1)k_1^2}{n\delta\left\{\frac{\delta^2K}{k_1d_1}-\frac{\delta}{n\kappa}-\frac{1}{n}\right\}}$$
$$\times\left\{\frac{3k_1d_2}{\delta}+(1+\frac{k_1}{\delta})\frac{3d_1L}{\delta}+\max\{\frac{g(0)}{\kappa},\frac{1}{\kappa}\}\right\}.$$

(Here, we used the fact that $\frac{\varphi}{n\varphi-1}$ decreases with φ and $\varphi\geq\frac{\delta^2K}{k_1d_1}-\frac{\delta}{n\kappa}$.)

Part (a) is now clear.

(b) It was shown in Piunovskiy (2009a) that the EOQ in the fluid model satisfies $\varphi^*\geq\frac{\delta^2K}{k_1d_1}$ and the EOQ in the (scaled) stochastic model ${}^n\Phi^*$ satisfies $\frac{{}^n\Phi^*}{n}\geq\frac{\delta^2K}{k_1d_1}-\frac{\delta}{n\kappa}$. Therefore, according to part (a), for $n\geq N$, we have

$${}^nTCU([n\varphi^*]) \;\leq\; tcu(\varphi^*)+E(n)\leq tcu(\frac{{}^n\Phi^*}{n})+E(n)\leq {}^nTCU({}^n\Phi^*)+2E(n)$$

in one direction; and

$${}^nTCU([n\varphi^*]) \;\geq\; {}^nTCU({}^n\Phi^*)\geq tcu(\frac{{}^n\Phi^*}{n})-E(n)\geq tcu(\varphi^*)-2E(n)$$

in the other direction. Combining both directions results in the statement. ■

Proof of Proposition 4.2. ${}^n\Pi(i)$ and $\pi(y)$ can be easily computed as done in Piunovskiy (2009a). So we have

$$\Pi(0)=\frac{\frac{1}{n\kappa}}{\sum_{j=1}^{[n\varphi]}\frac{1}{n\mu(\frac{j}{n})}+\frac{1}{n\kappa}};\ \Pi(i)=\frac{\frac{1}{n\mu(i/n)}}{\sum_{j=1}^{[n\varphi]}\frac{1}{n\mu(\frac{j}{n})}+\frac{1}{n\kappa}},\ i=1,2,\ldots,[n\varphi];$$

and

$$\pi(0)=0;\ \pi(y)=\frac{1}{\mu(y)t_{cycle}},\ 0<y\leq[n\varphi].$$

Then $|{}^n\Pi(0)-\pi(0)|\leq\frac{1}{n\kappa(\frac{1}{n\kappa}+\frac{[n\varphi]}{nk_1})}\leq\frac{k_1}{k_1+\kappa[n\varphi]}$; and

$$|{}^n\Pi(i)-\int_{\frac{i-1}{n}}^{\frac{i}{n}}\pi(y)dy| \;=\; \left|\frac{\frac{1}{n\mu(i/n)}}{E_{[n\varphi]}[{}^nT_{cycle}]}-\frac{\int_{\frac{i-1}{n}}^{\frac{i}{n}}\frac{1}{\mu(y)}dy}{t_{cycle}}\right|$$

$$= \left| \frac{1}{n\mu(i/n)} t_{cycle} - \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{\mu(y)} dy\, t_{cycle} + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{\mu(y)} dy\, t_{cycle} - \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{\mu(y)} dy E_{[n\varphi]}[{}^nT_{cycle}] \right| \frac{1}{E_{[n\varphi]}[{}^nT_{cycle}]t_{cycle}}$$

$$\leq \frac{t_{cycle}\left|\frac{1}{n\mu(i/n)} - \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{dy}{\mu(y)}\right| + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{dy}{\mu(y)} \left|E_{[n\varphi]}[{}^nT_{cycle}] - t_{cycle}\right|}{E_{[n\varphi]}[{}^nT_{cycle}]t_{cycle}}$$

$$\leq \frac{\frac{2\varphi}{n\delta^2} + \frac{1}{n\delta}\left(\frac{B_1}{n} + B_2\tilde{\eta}^{-n2[\varphi+1]} + \frac{1}{n\kappa}\right)}{\left(\frac{[n\varphi]}{nk_1} + \frac{1}{n\kappa}\right)\frac{\varphi}{k_1}}.$$

(Here we recall part (a) of Lemma 4.3.)

Recall that in the above derived expression, $\tilde{\eta}$ can be any number from $(1,\infty)$. After passing to the limit $\tilde{\eta} \to \infty$, we eventually end up with

$$|{}^n\Pi(i) - \int_{\frac{i-1}{n}}^{\frac{i}{n}} \pi(y)dy| \leq \frac{\frac{2\varphi k_1^2\kappa}{\delta^2} + \frac{k_1^2\kappa}{n\delta}\left(\frac{3k_1d_2(\varphi+1)}{\delta} + (1+\frac{k_1}{\delta})\frac{3d_1L}{\delta} + \frac{1}{\kappa}\right)}{([n\varphi]+1)\varphi},$$

as required. ■

4.6.2 Proof of Proposition 4.3, Corollary 4.3 and Proposition 4.4

Let us call a cycle the time duration between two consecutive moments when the inventory is backed up full. Arguing similarly as in the beginning of Subsection 4.6.1, it suffices to consider the inventory level process $\{{}^nY_t,\ t \geq 0\}$ and the cost incurred over one complete cycle, for which we put the starting time of $t = 0$. Let us denote t_{cycle}, ${}^nT_{cycle}$ and tc, ${}^nTC$ the duration of a cycle and the cost incurred over a cycle in the fluid and scaled stochastic model, respectively. Notice additionally that a cycle is always constituted to by two phases corresponding to the on and off of the production. This raises another set of denotations: let t_{on}, ${}^nT_{on}$ (t_{off}, ${}^nT_{off}$) and tc_{on}, ${}^nTC_{on}$ (tc_{off}, ${}^nTC_{off}$) be the total cost incurred during the production-on (off) phases in the fluid and (scaled) stochastic model, respectively. We agree on that in both fluid and (scaled) stochastic model, the setup cost is accounted for in tc_{off} and ${}^nTC_{off}$. Then obviously we have $t_{cycle} = t_{on} + t_{off}$, $E[{}^nT_{cycle}] = E[{}^nT_{on}] + E[{}^nT_{off}]$ and $tc = tc_{on} + tc_{off}$, $E[{}^nTC] = E[{}^nTC_{on}] + E[{}^nTC_{off}]$. Here and below, for convenience we have omitted the subscript of the expectation operator.

Lemma 4.4 *Under Conditions 4.2, the following inequalities hold:*
(a) $|E[{}^nTC_{on}] - tc_{on}| \leq \frac{B_1}{n} + B_2\tilde{\eta}^{-2n[\varphi+1]}$, $|E[{}^nT_{on}] - t_{on}| \leq \frac{B_1}{n} + B_2\tilde{\eta}^{-2n[\varphi+1]}$;
(b) With nonnegative B_1 *and* B_2 *given in (4.3) and (4.4):*

$$\left|E[{}^nTC_{off}] - tc_{off}\right| \leq \frac{B_1}{n} + B_2\tilde{\eta}^{-2n[\varphi+1]} + \frac{g(0)}{n\kappa},$$

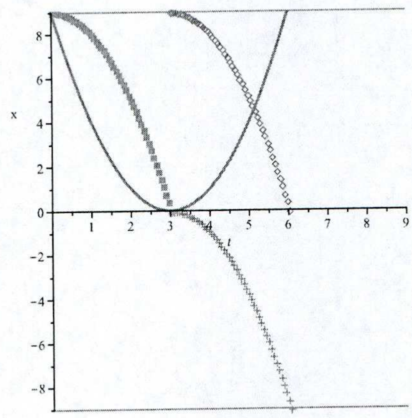

Figure 4.1: The illustrative graph of ${}_{on}\hat{y}(t)$.

$$\left|E[{}^{n}T_{off}]-t_{off}\right| \quad \le \quad \frac{B_1}{n}+B_2\tilde{\eta}^{-2n[\varphi+1]}+\frac{1}{n\kappa}.$$

Proof. (a) Let us concentrate on the inventory level process over the production-on phase. In the fluid model, it appears convenient to reflect the trajectory $\{y(t), t \in (t_{on}, t_{cycle})\}$ (corresponding to the solid curve in Figure 4.1) about the horizontal t-axis first, and then shift the resulting trajectory (corresponding to the curve of crosses in Figure 4.1) upwards by φ units, and finally further shift the resulting trajectory to the left by shifting the time by t_{off} units to the right to get $\{{}_{on}\hat{y}(t), t \in [0, t_{on}]\}$ (corresponding to the curve of solid boxes in Figure 4.1). Note now, for $\{{}_{on}\hat{y}(t), t \in [0, t_{on}]\}$ with ${}_{on}\hat{y}(0) = \varphi$ the roles of production and demand have switched over: each produced unit reduces ${}_{on}\hat{y}$ by one unit, and each demanded unit increases ${}_{on}\hat{y}$ by one unit. More precisely, let us define the following functions

$$\begin{aligned}
&{}_{on}\hat{\mu}(\varphi) = \lambda(0), {}_{on}\hat{\mu}(0) = 0,\ {}_{on}\hat{\mu}(y) = \lambda(\varphi - y),\ y \in (0,\varphi);\\
&{}_{on}\hat{\lambda}(\varphi) = \mu(0) = 0,\ {}_{on}\hat{\lambda}(0) = 0,\ {}_{on}\hat{\lambda} = \mu(\varphi - y),\ y \in (0,\varphi);\\
&{}_{on}\hat{g}(\varphi) = g(0),\ {}_{on}\hat{g}(0) = 0,\ {}_{on}\hat{g}(y) = g(\varphi - y),\ y \in (0,\varphi),
\end{aligned}$$

so that the dynamics of ${}_{on}\hat{y}(0) = \varphi$ with $\left.\frac{d\,{}_{on}\hat{y}(t)}{dt}\right|_{{}_{on}\hat{y}(t)=y} = {}_{on}\hat{\lambda}(y) - {}_{on}\hat{\mu}(y)$ for $y \in (0,\varphi]$, where we take the left derivative for the case of ${}_{on}\hat{y} = \varphi$, is of our interest, because now we can write $tc_{on} = \int_0^\infty g({}_{on}\hat{y}(t))dt$.

Absolutely similar arguments are applicable to the (scaled) stochastic model. Consequently, we can consider the inventory level process during a production-on phase as a Birth-and-Death process $\{{}^{n}_{on}\hat{Y}_t, t \ge 0\}$ with initial condition ${}^{n}_{on}\hat{Y}_0 = [n\varphi]$, state space $\{0, 1, \ldots\}$, birth and death rates given by $n\alpha(\frac{i}{n}) \triangleq n\,{}_{on}\hat{\lambda}(\frac{i}{n})$ and $n\beta(\frac{i}{n}) \triangleq n\,{}_{on}\hat{\mu}(\frac{i}{n})$

when the current state is i, and the cost rate given by $\gamma(\frac{i}{n}) \triangleq {}_{on}\hat{g}(\frac{i}{n})$. Now, recognizing $E_{[n\varphi]}[\int_0^\infty {}_{on}\hat{g}({}^n_{on}\hat{Y}_t)dt] = E[{}^nTC_{on}]$ and that Conditions 4.2 is a special version of Conditions 4.3, we can refer to Lemma 4.2 for $|{}^nE[{}^nTC_{on}] - tc_{on}| \leq \frac{B_1}{n} + B_2\tilde{\eta}^{-2n[\varphi+1]}$. Arguing similarly as above (see also the proof of Lemma 4.3), we have $|{}^nE[{}^nT_{on}] - t_{on}| \leq \frac{B_1}{n} + B_2\tilde{\eta}^{-2n[\varphi+1]}$. Part (a) is now clear.

(b) The production-off phase has already been covered when analyzing EOQ models. Therefore, one can directly refer to Lemma 4.3 for the statement. ■

Proof of Proposition 4.3. Lemma 4.4 implies that

$$\begin{aligned}
|E[{}^nTC] - tc| &\leq |E[{}^nTC_{on}] - tc_{on}| + |E[{}^nTC_{off}] - tc_{off}| \\
&\leq \frac{2}{n}\left[B_1 + B_2 n\tilde{\eta}^{-2n[\varphi+1]} + \frac{g(0)}{2\kappa}\right]
\end{aligned}$$

and similarly $|E[{}^nT_{cycle}] - t_{cycle}| \leq \frac{2}{n}\left[B_1 + B_2 n\tilde{\eta}^{-2n[\varphi+1]} + \frac{1}{2\kappa}\right]$. Then according to (4.5) and the facts of $E[{}^nT_{cycle}] \geq \frac{2[n\varphi]}{nk_1} + \frac{1}{n\kappa}$, $\frac{2\varphi}{\bar{\delta}} \geq t_{cycle} \geq \frac{2\varphi}{\bar{\delta}}$ and $tc \leq \frac{2\varphi d_1}{\underline{\delta}}$ (recall $\bar{\delta} \triangleq \max\{\delta, \delta_{\mu\lambda}\}$ and $\underline{\delta} \triangleq \min\{\delta, \delta_{\mu\lambda}\}$), we have

$$\begin{aligned}
&|{}^nTCU([n\varphi]) - tcu(\varphi)| \\
\leq\ & \frac{\frac{2\varphi}{\underline{\delta}}\frac{2}{n}\left\{B_1 + B_2 n\tilde{\eta}^{-2n[\varphi+1]} + \max\{\frac{g(0)}{2\kappa}, \frac{1}{2\kappa}\}\right\}(1+d_1)}{\left(\frac{2[n\varphi]}{nk_1} + \frac{1}{n\kappa}\right)\frac{2\varphi}{\bar{\delta}}} \\
=\ & \frac{2\bar{\delta}\kappa k_1(1+d_1)}{\underline{\delta}}\left\{\frac{B_1 + B_2 n\tilde{\eta}^{-2n[\varphi+1]} + \max\{\frac{g(0)}{2\kappa}, \frac{1}{2\kappa}\}}{2[n\varphi]\kappa + k_1}\right\}.
\end{aligned}$$

■

Proof of Corollary 4.3. (a) Suppose the statement does not hold. That is, for some subsequence $\{n_j, j = 1, 2, \ldots\}$ with $n_j \to \infty$ as $j \to \infty$, ${}^{n_j}\Phi^* = o(n_j)$ in that $\lim_{j\to\infty} \frac{{}^{n_j}\Phi^*}{n_j} = 0$. Now fixing ${}^{n_j}\Phi^*$, we have

$$\begin{aligned}
E[{}^{n_j}T_{cycle}] &= \left\{\frac{1}{n_j\kappa} + \sum_{i=1}^{{}^{n_j}\Phi^*}\frac{1}{n_j\mu(\frac{i}{n_j})}\right\} \\
&+ \sum_{i=0}^{{}^{n_j}\Phi^*-1}\left\{\frac{1}{n_j\lambda(\frac{i}{n_j})} + \sum_{k=0}^{i-1}\frac{\mu(\frac{i}{n_j})\mu(\frac{i-1}{n_j})\ldots\mu(\frac{i-k}{n_j})}{n_j\lambda(\frac{i}{n_j})\lambda(\frac{i-1}{n_j})\ldots\lambda(\frac{i-k-1}{n_j})}\right\},
\end{aligned}$$

where by (Wang and Yang 2005, Thm.1, p.175) the term inside the first curry bracket corresponds to $E[{}^{n_j}T_{off}]$ and the second (last) sum corresponds to $E[{}^{n_j}T_{on}]$. Here, we agree on that when $i = 0$, the term in the second curry bracket reduces to $\frac{1}{n_j\lambda(0)}$. This gives

$$
\begin{aligned}
& {}^{n_j}TCU({}^{n_j}\Phi^*) \\
= & \frac{E[{}^{n_j}TC]}{\left\{\frac{1}{n_j\kappa}+\sum_{i=1}^{{}^{n_j}\Phi^*}\frac{1}{n_j\mu(\frac{i}{n_j})}\right\}+\sum_{i=0}^{{}^{n_j}\Phi^*-1}\left\{\frac{1}{n_j\lambda(\frac{i}{n_j})}+\sum_{k=0}^{i-1}\frac{\mu(\frac{i}{n_j})\mu(\frac{i-1}{n_j})\ldots\mu(\frac{i-k}{n_j})}{n_j\lambda(\frac{i}{n_j})\lambda(\frac{i-1}{n_j})\ldots\lambda(\frac{i-k-1}{n_j})}\right\}} \\
\geq & \frac{K}{\left\{\frac{1}{n_j\kappa}+\sum_{i=1}^{{}^{n_j}\Phi^*}\frac{1}{n_j\mu(\frac{i}{n_j})}\right\}+\sum_{i=0}^{{}^{n_j}\Phi^*-1}\left\{\frac{1}{n_j\lambda(\frac{i}{n_j})}+\sum_{k=0}^{i-1}\frac{\mu(\frac{i}{n_j})\mu(\frac{i-1}{n_j})\ldots\mu(\frac{i-k}{n_j})}{n_j\lambda(\frac{i}{n_j})\lambda(\frac{i-1}{n_j})\ldots\lambda(\frac{i-k-1}{n_j})}\right\}},
\end{aligned}
$$

where the last inequality follows from the fact $g(y) \geq 0$. Clearly, the right hand side expression of the above inequality goes to ∞ as $n_j \to \infty$, because $\lambda(y)$ and $\mu(y)$ are both bounded and separated from zero, and ${}^{n_j}\Phi^* = o(n_j)$ by supposition. On the other hand, obviously there exists some $\varphi_* > 0$ with $tcu(\varphi_*) < \infty$, which according to Proposition 4.3 leads to that at least for big enough n_j, $[n_j\varphi_*]$ outperforms ${}^{n_j}\Phi^*$, which is a desired contradiction. Part (a) is thus proved.

(b) The proof of this part is the same as that of Corollary 4.1 and thus omitted.

(c1) Let us notice first of all that under the conditions in the statement, we have $\frac{\tilde{B}_1+\tilde{B}_2 n\tilde{\eta}^{-2n(\varphi+1)}+\max\{\frac{g(0)}{2\kappa},\frac{1}{2\kappa}\}}{2(n\varphi-1)\kappa+k_1} > 0$ and decreases with φ. Here, $\tilde{B}_1$ and $\tilde{B}_2$ come from replacing $[\varphi+1]$ by $\varphi+1$ in B_1 and B_2. Indeed, as for the positivity part, one only needs to see the denominator $2(n\varphi-1)\kappa+k_1 > 0$ if n is subject to the given condition. The decreasing (with respect to φ) part follows from $\frac{d[\frac{\varphi+1}{2\kappa(n\varphi-1)+k_1}]}{d\varphi} = \frac{k_2-2\kappa-2\kappa n}{[2\kappa(n\varphi-1)+k_1]^2} < 0$ whenever $n > \frac{k_1}{2\kappa}-1$. Remember, $\tilde{B}_1$ and $\tilde{B}_2$ are both φ-dependent, and L is φ-independent here.

Now let us prove part (c1) of the corollary. Noticing that under the conditions for the statement, Proposition 4.3 implies that

$$
\begin{aligned}
& |{}^{n}TCU([n\varphi]) - tcu(\varphi)| \\
\leq & \frac{2\tilde{\delta}\kappa k_1(1+d_1)}{\underline{\delta}}\left\{\frac{\tilde{B}_1+\tilde{B}_2 n\tilde{\eta}^{-2n(\varphi+1)}+\max\{\frac{g(0)}{2\kappa},\frac{1}{2\kappa}\}}{2(n\varphi-1)\kappa+k_1}\right\}.
\end{aligned}
\tag{4.6}
$$

For $\varphi \geq \frac{K\delta^2}{4k_1d_1}$, one can bound from the above the right hand side expression in (4.6) by substituting $\varphi = \frac{K\delta^2}{4k_1d_1}$ in it, which leads to the part (c1).

(c2) Let us notice that $\varphi^* \geq \frac{K\delta^2}{4k_1d_1}$, and for big enough n, $\frac{{}^{n}\Phi^*}{n} \geq \frac{K\delta^2}{4k_1d_1}$. Indeed, due to part (b), to justify the second inequality, we only need verify $\varphi^* \geq \frac{K\delta^2}{4k_1d_1}$, which is done as follows. For the fluid model, clearly we have

$$
tcu(\varphi) = \frac{\int_0^{\varphi}\frac{g(y)}{\mu(y)}dy+\int_0^{\varphi}\frac{g(y)}{\lambda(y)-\mu(y)}+K}{\int_0^{\varphi}\frac{dy}{\mu(y)}+\int_0^{\varphi}\frac{dy}{\lambda(y)-\mu(y)}},
$$

where the numerator corresponds to tc and the denominator to t_{cycle}, and

$$\frac{d\,tcu}{d\varphi} = \frac{(\frac{g(\varphi)}{\mu(\varphi)} + \frac{g(\varphi)}{\lambda(\varphi)-\mu(\varphi)})[\int_0^\varphi \frac{dy}{\mu(y)} + \int_0^\varphi \frac{dy}{\lambda(y)-\mu(y)}]}{[\int_0^\varphi \frac{dy}{\mu(y)} + \int_0^\varphi \frac{dy}{\lambda(y)-\mu(y)}]^2} - \frac{[\int_0^\varphi \frac{g(y)}{\mu(y)}dy + \int_0^\varphi \frac{g(y)}{\lambda(y)-\mu(y)} + K](\frac{1}{\mu(\varphi)} + \frac{1}{\lambda(\varphi)-\mu(\varphi)})}{[\int_0^\varphi \frac{dy}{\mu(y)} + \int_0^\varphi \frac{dy}{\lambda(y)-\mu(y)}]^2},$$

which is negative if $(\frac{g(\varphi)}{\mu(\varphi)} + \frac{g(\varphi)}{\lambda(\varphi)-\mu(\varphi)})[\int_0^\varphi \frac{dy}{\mu(y)} + \int_0^\varphi \frac{dy}{\lambda(y)-\mu(y)}] < \frac{K}{\mu(t)}$. But the latter inequality holds if $\varphi < \frac{K\delta^2}{4k_1d_1}$, because

$$\frac{g(\varphi)}{\mu(\varphi)} + \frac{g(\varphi)}{\lambda(\varphi)-\mu(\varphi)} \le \frac{2d_1}{\underline{\delta}}$$

and

$$\int_0^\varphi \frac{dy}{\mu(y)} + \int_0^\varphi \frac{dy}{\lambda(y)-\mu(y)} \le \frac{2\varphi}{\underline{\delta}}.$$

This means that EOQ in the fluid model satisfies $\varphi^* \ge \frac{K\delta^2}{4k_1d_1}$.

Now with the help of (c1) and Proposition 4.3, part (c2) can be proved in the same way as for part (b) of Corollary 4.2. ■

Proof of Proposition 4.4. For brevity, in this proof we shall denote "on" by a bar, so that $\overline{0}$ is read as $(0, on)$. In addition, we shall view $\overline{i}$ the same as i, as far as arithmetics are concerned: for example, $2^{\overline{i}} = 2^i$.

It is well known (see Zhu and Prieto-Rumeau (2008) as well as (Piunovskiy 2009a, Proof of Thm.3)) that $''TCU(Z)$ satisfies the following equations: $\forall i \in \{1, 2, \ldots, Z\}$,

$$h\frac{i}{n} + KnDI\{i=1\} + nDV(i-1) - nDV(i) = {''TCU(Z)}; \tag{4.7}$$

$\forall i \in \{\overline{1}, \ldots \overline{Z-1}, Z\}$,

$$\frac{h}{n}i + nDV(i-1) + nRV(i+1) - (nR+nD)V(i) = {''TCU(Z)}; \tag{4.8}$$

and at $i = \overline{0}$,

$$\begin{aligned} nRV(\overline{1}) - nRV(\overline{0}) &= {''TCU(Z)} \\ \Leftrightarrow \quad V(\overline{0}) &= V(\overline{1}) - \frac{''TCU(Z)}{nR}. \end{aligned} \tag{4.9}$$

From (4.7) we have

$$V(i) = K + V(\overline{0}) + \frac{\frac{h\sum_{j=1}^{j=i} j}{n} - i\,{''TCU(Z)}}{nD}. \tag{4.10}$$

As for (4.8), according to (Piunovskiy 2004a, p.169), its homogeneous version has a general solution $V_h(i) = C_1 + C_2\left(\frac{D}{R}\right)^i$, where C_1, C_2 are some constants. Furthermore, one can check that a particular solution to (4.8) is $V_p(i) = \frac{hi}{n^2(D-R)}\left[\frac{i+1}{2} + \frac{R}{D-R}\right] - \frac{"TCU(Z)i}{n(D-R)}$, which thus leads to the general solution in the form

$$V(i) = C_1 + C_2\left(\frac{D}{R}\right)^i + \frac{hi}{n^2(D-R)}\left[\frac{i+1}{2} + \frac{R}{D-R}\right] - \frac{"TCU(Z)i}{n(D-R)}. \tag{4.11}$$

According to (4.11) and (4.9), we have $C_1 = V(\overline{1}) - \frac{"TCU(Z)}{nR} - C_2$.
At $i = \overline{1}$, we have

$$\begin{aligned} V(\overline{1}) = V(\overline{1}) - \frac{"TCU(Z)}{nR} - C_2 + C_2\left(\frac{D}{R}\right)^{\overline{1}} + \frac{h\overline{1}}{n^2(D-R)}\left[\frac{\overline{1}+1}{2} + \frac{R}{D-R}\right] \\ - \frac{"TCU(Z)\overline{1}}{n(D-R)} \end{aligned}$$

$$\begin{aligned} \Leftrightarrow\ & 0 = -\frac{"TCU(Z)}{nR} + C_2\frac{D-R}{R} + \frac{h}{(D-R)n^2}\frac{D}{D-R} - \frac{"TCU(Z)}{(D-R)n} \\ \Leftrightarrow\ & C_2\frac{D-R}{R} = \frac{"TCU(Z)}{n(D-R)} + \frac{"TCU(Z)}{nR} - \frac{h}{(D-R)n^2}\frac{D}{D-R} \\ \Leftrightarrow\ & C_2 = \frac{"TCU(Z)\,R}{n(D-R)^2} + \frac{"TCU(Z)}{n(D-R)} - \frac{RhD}{n^2(D-R)^3}. \end{aligned}$$

Due to the "continuity", at $i = Z$, $V(i)$ given by (4.10) and (4.11) must coincide. That is,

$$\begin{aligned} \frac{h(1+Z)Z - 2Zn\,"TCU(Z)}{2n^2D} + K \ = \ & \left[\frac{"TCU(Z)\,R}{n(D-R)^2} + \frac{"TCU(Z)}{n(D-R)}\right. \\ & \left. - \frac{RDh}{n^2(D-R)^3}\right] \times \left[\left(\frac{D}{R}\right)^Z - 1\right] \\ & + \frac{hZ}{n^2(D-R)}\left[\frac{Z+1}{2} + \frac{R}{D-R}\right] \\ & - \frac{"TCU(Z)Z}{n(D-R)}, \end{aligned}$$

resulting in

$$\begin{aligned} "TCU(Z) \ = \ & \{2hZ^2DR^2 - hZ^2D^2R - 3hZD^2R + 4hZDR^2 - 2RhD^2 \\ & -hZR^3 - hZ^2R^3 + 2Kn^2D^4 - 6Kn^2D^3R + 6Kn^2D^2R^2 \\ & -2Kn^2DR^3 + 2RhD^2(\frac{D}{R})^Z\Big\}\frac{1}{2n} \\ & \times\frac{1}{D^2R - ZR^3 - D^3 - ZRD^2 + 2ZR^2D + (D^3 - RD^2)(\frac{D}{R})^Z}. \end{aligned}$$

Now it only remains to factorize and simplify the above expression. ■

Chapter 5

Fluid approximation: BS networks

5.1 Introduction

In Chapter 3 and Chapter 4, we studied a state-dependent translation of a fluid optimal policy into one for the stochastic model, and showed its efficiency is of order $\frac{1}{n}$, where n is the fluid scaling parameter. In this chapter by considering a bandwidth-sharing network, we study the question: do different translations of fluid optimal policies provide qualitatively different efficiencies?

A bandwidth-sharing network can be described as follows. Through a network of J resources (links), L flows are routed in a predetermined way. The L flows are assumed to be distinct in the sense that each one is routed differently. The flows represent aggregate streams of files of finite size. Thus the files are classified according to which flow they belong. Files belonging to flow l are called files of type l. Each resource has a finite capacity shared among the flows passing through it. Let us introduce some notations. A network is described by a configuration (R,Z), where R is a $J \times L$ matrix with $r_{jl} = 1$ if resource j participates in serving files of type l and $r_{jl} = 0$ otherwise. We assume that each column of R has at least one nonzero element, meaning that each file (flow) must be served somewhere. Here unless stated otherwise, by a vector we always mean a column vector. Let Z be a J vector with $z_j > 0$ indicating the maximal capacity of resource j. Let $\Phi \geq 0$, an L vector, which can be dependent on current time as well as the number of flows, represent instantaneous allocations of resources for each type of files. The vector Φ is subject to the constraint $R\Phi \leq Z$, where $\leq$ is understood to be componentwise. To our best knowledge, this bandwidth-sharing network model was originated in Massoulié and Roberts (2000). Later it has been extensively studied. An interested reader can find a thorough literature review about the model in Verloop

(2009).

We shall model the network dynamics as a Markovian system. In greater details, denoting by ${}^nY_t$, an L vector, the instantaneous number of files of each type in the network, we suppose that $\{{}^nY_t,\ t \geq 0\}$ is a continuous time Markov chain, with $S = \mathbb{Z}_+^L$ the state space, $n\lambda_l$ the transition rate from ${}^nY$ to ${}^nY + e_l$ and $n\mu_l\Phi_l$ the transition rate from ${}^nY$ to ${}^nY - e_l$ (if the corresponding element of ${}^nY$ is positive). Here and below we agree on that $\mathbb{R}^L$ ($\mathbb{R}_+^L$, $\mathbb{Z}_+^L$) indicates the L vectors of real (nonnegative real, nonnegative integer) numbers, and denote by $\{e_l\}$ the natural basis of $\mathbb{R}^L$. This is saying that files of type l arrive in a Poisson process with intensity $n\lambda_l$, and each file is of size exponentially distributed with mean $\frac{1}{n\mu_l}$, and hence the mean time for its service to be completed will be the ratio of Φ_l, its allocated capacity, against $\frac{1}{n\mu_l}$. Suppose the performance measure of the system is the total holding cost in the form of

$$E^{\Phi}_{{}^nY_0}\left[\int_0^T e'\frac{{}^nY_t}{n}dt\right],$$

where $E^{\Phi}_{{}^nY_0}$ is the expectation with an initial state ${}^nY_0$ and resource allocation Φ, $T > 0$ is a finite horizon, $e \triangleq (1,\dots,1)'$ and $'$ stands for the transposition. Here as in Chapter 3 and Chapter 4 $n \in \mathbb{N}$ is the scaling parameter. Then one aims at finding the optimal resource allocation to minimize this performance functional. The optimization with this criterion can be interpreted as the minimization of the total workload. More discussions about this criterion can be found in Verloop (2009).

On the other hand, if we let $y(t)$ and u be the analogues of ${}^nY_t$ and Φ, and further agree on the following notations: $\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_L)'$,

$$M = diag(\mu_1, \mu_2, \dots, \mu_L) \triangleq \begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \mu_2 & \dots & 0 \\ 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \mu_L \end{pmatrix},$$

then the fluid model to the above described bandwidth-sharing network can be written down as the following linear program:

$$\begin{aligned} & \int_0^T e'y(t)dt \to \min_{u,y} \qquad (5.1) \\ \text{s.t.} \quad & \frac{dy}{dt} = -Mu(t) + \Lambda, \\ & y(0) = Y_0 \triangleq \frac{{}^nY_0}{n}, \\ & Ru(t) \leq Z, \\ & y(t), u(t) \geq 0. \end{aligned}$$

Note that here the state is continuous, and the state space is $\mathbb{R}_+^L$.

Suppose now a fluid optimal policy $u^*(t)$ is obtained, whose existence is guaranteed by Lemma 5.1 below, one may have many options in translating it into one for the stochastic model with a satisfactory performance. For example, in Chapter 3 and Chapter 4, we considered a state-dependent translation. The policy resulted in by that translation requires to know the exact state, and thus more memory about the system. In this relation, a cheaper translation would be via the following translation:

$$\Phi_l^*({}^nY_t^l, t) = u_l^*(t) I\{{}^nY_t^l > 0\},$$

$l = 1, 2, \ldots, L$, where u_l^*, Φ_l^* and ${}^nY_t^l$ are the lth element of u^*, Φ^*, and ${}^nY_t$. Clearly, the resulting policy, the so called tracking policy (Bäuerle (2000)), is nearly state-independent in that the only matter is whether the state reaches zero or not, instead of the exact state. Then one is interested in comparing the following quantities:

$${}^n\hat{W}({}^nY_0) \triangleq E_{{}^nY_0}^{\Phi^*}\left[\int_0^T e' \frac{{}^nY_t}{n} dt\right],$$

$$\hat{w}(Y_0) \triangleq \int_0^T e' y(t) dt,$$

where $\hat{w}(Y_0)$ is the performance functional for the fluid model under its optimal control. To be rigorous we should have indicated the control policy. But we omit that, since under a fixed policy, we actually deal with essentially uncontrolled processes. Since ${}^nY_0$ and Y_0 are fixed, from now on we shall even only write ${}^n\hat{W}$ and $\hat{w}$ instead of ${}^n\hat{W}({}^nY_0)$ and $\hat{w}(Y_0)$ for notational convenience.

Revolving about $\left|{}^n\hat{W} - \hat{w}\right|$, the following definition is in position: call the tracking policy Φ^* AFO (Gajrat and Hordijk (2000); Gajrat et al (2003); Maglaras (2000)) if

$$\lim_{n \to \infty} \left|{}^n\hat{W} - \hat{w}\right| = 0.$$

The rate of convergence is the efficiency of the underlying translation. It should be noted that different translations could all result in AFO policies. Therefore, their efficiencies can be used as a measure when comparing them.

To our best knowledge, the above described tracking policy was initially studied in Bäuerle (2000), where the author considered a discounted scheduling problem for a multi-class queueing network. In this chapter, for the aforementioned bandwidth-sharing network on a finite horizon, we shall address the following two questions:

- Is tracking policy AFO?
- In case of an affirmative answer to the first question, how efficient is the tracking policy?

It should be noted that for the problem considered in Bäuerle (2000), the author did not study the second question, which as mentioned in the previous paragraph is indeed important. To our best knowledge, the question about efficiency of the tracking policy was only very briefly mentioned in Gajrat et al (2003) for a special discrete time tandem queueing system.

In the current chapter, for the concerned finite horizon problem, we firstly answer the first question affirmatively. In the meanwhile, we provide two answers to the second question with upper boundary estimates for the rate of convergence. Specifically, we show that the tracking policy can be efficient or much less depending on the parameters of the model. While one example was mentioned in Gajrat et al (2003) to say that tracking policy is less efficient than a feedback (state-dependent) translation considered, in this chapter we shall provide an opposite example to show that tracking policy can be also efficient, and hence favored due to the less information it requires to be implemented.

The rest of this chapter is organized as follows. In Section 5.2 we state the main results, which are verified with a simple but illuminative example in Section 5.3. Finally we sum up this chapter with a conclusion. The sketched proofs of the main statements alongside auxiliary lemmas are included in Section 5.5 in the very end of this chapter.

5.2 Main statements

In this section we shall provide two answers to the question about the efficiency of tracking policy raised at the end of Section 5.1. The first answer, given by Theorem 5.1, is corresponding to the general case where we do not impose extra conditions on the parameters of the network; while the second answer, given in Theorem 5.2, looks more interesting, but is based on some extra conditions.

Let us start with a lemma, giving the form of the fluid optimal policy, following which the main denotations used in this section, in addition to those introduced in section 5.1, can be introduced.

Lemma 5.1 *There is an optimal resource allocation for the fluid model (5.1), and the optimal resource allocation is a piecewise constant function in time, namely, $u^*(t)$, with N subintervals $[T_i, T_{i+1})$, where $N \in \mathbb{N}$, $i = 0, 1, \ldots, N-1$, $T_0 \triangleq 0$ and $T_N \triangleq T$.*

Consider files of type l (thus the lth "queue") by subintervals, so that we may use the language from Queueing theory. Let ${}^i\mu_l \triangleq \mu_l u^*(t)$, $i = 1, \ldots, N$ be the potential service rate of the lth queue on the ith subinterval.

5.2.1 Efficiency of the tracking policy for the general case

Now we are in the position to state the following:

Theorem 5.1 *Let us put* $\gamma \triangleq \max_{i=1,\dots,N,l=1,\dots,L}\{\lambda_l, {}^i\mu_l\}$ *and* $\bar{T} \triangleq \max_{i=1,\dots,N}\{T_i - T_{i-1}\}$. *For any fixed initial state and network configuration, the above quantities are fixed. Then*

$$|{}^n\hat{W} - \hat{w}| \leq \frac{TL}{\sqrt{n}}\left\{12\sqrt{N(N+1)\gamma\bar{T}}\right\}.$$

It will be confirmed by an example in Section 5.3 that as far as the order is concerned, the obtained $\frac{1}{\sqrt{n}}$ is correct and cannot be improved. It should be noted that Theorem 5.1 does not require extra conditions on the primary data. Then with Theorem 5.1, we say that the efficiency of the tracking policy is $O(\frac{1}{\sqrt{n}})$. It is then interesting to compare the efficiency of the tracking policy with that of the feedback translation, as considered in for example, Chapter 3. There, for a controlled $M/M/1$ queueing system with state zero as the single absorbing state, we considered a feedback translation, which was shown to be of efficiency $O(\frac{1}{n})$. If we consider a corresponding bandwidth-sharing network with only one resource and flow, then Theorem 5.1 somehow suggests that the tracking policy is in general less efficient than feedback type translations. This is in line with the result in Gajrat et al (2003), where for a different model the author also observed that the tracking policy could be less efficient than feedback translations.

On the other hand, suppose we still consider the simple case of $M/M/1$ queue (that is, the bandwidth-sharing network of one resource and one flow). In addition, we impose some conditions on the parameters so that for the given $Y_0 > 0$, the time horizon T is sufficiently small. Then compared to the absorbing case, we should expect that the convergence $|{}^n\hat{W} - \hat{w}| \to 0$ should go not more slowly than $O(\frac{1}{n})$, because we now count the deviation of the stochastic model from the fluid model for less time. This motivates our study of a preabsorbing case in the next subsection, where the time horizon is small enough so that over it for each type of files, the fluid model does not reach zero.

5.2.2 Efficiency of the tracking policy for a preabsorbing case

Conditions 5.1 *Uniformly in l, T is such that* $Y_0^l - |\underline{\lambda} - \bar{\mu}|T \triangleq \underline{y} > 0$ *and* $Y_0^l + \bar{\lambda}T \triangleq \bar{y} < Y$. *Here* $Y > 0$, $\underline{\lambda} \triangleq \min_{l=1,\dots,L}\{\lambda_l\}$, $\bar{\lambda} \triangleq \max_{l=1,\dots,L}\{\lambda_l\}$ *and* $\bar{\mu} \triangleq \max_{l=1,\dots,L}\{\mu_l\} \times \max_{l=1,\dots,L}\{z_l\}$.

We call a bandwidth-sharing network preabsorbing if Conditions 5.1 are satisfied. Here, "preabsorbing" emphasizes that the time horizon is relatively short so that in the fluid model for each type of files, the "absorbing" state zero is not going to be attained. We note that $Y > 0$ in Conditions 5.1 can always be fixed, whatever the parameters of the network are.

Theorem 5.2 *Suppose Conditions 5.1 are satisfied. Then there exist L positive func-*

tions $\eta^l(n)$, each of which converges to zero as $n \to \infty$ such that

$$|^n\hat{W} - \hat{w}| \leq \sum_{l=1}^{L} \eta^l(n) = C \cdot O(\frac{1}{n}).$$

Here, the exact expressions for the functions η^l can be determined from the primary data by scanning the proof of the theorem.

Theorem 5.2 implies that the efficiency of the tracking policy can be much improved with the parameters of the model. We shall give an intuitive explanation for this improvement at the end of Section 5.3.

5.3 Example

To illustrate the obtained result, let us consider the following example, where $L = J = 1$. Suppose $\mu_1 = 2\mu > 0$, $\lambda_1 = \mu$, $z_1 = 1$, and that the initial system is empty, $^nY_0 = ny(0) = 0$. We shall come back to the reasons for this setting at the end of this section. Clearly, the fluid optimal policy is $u^*(t) = \frac{1}{2}$ on $[0,T]$, so that $\int_0^T y^*(t)dt = 0$, where $y^*(t)$ indicates the system under the optimal control. Now let us translate the fluid optimal policy into the policy for the scaled stochastic model and apply it. Then for any fixed scaling parameter n, we effectively deal with an M/M/1 queue $\{^nY_t, t \geq 0\}$ with arrival and service rates both equal to $n\mu$. We aim to compute $|E_0^{\Phi^*}\left[\int_0^T \frac{^nY_t}{n}dt\right]|$. Below we shall often omit the superscript and subscript for brevity. Since the trajectory in the fluid model will stay at zero, this provides the actual rate of convergence. Then we shall compare it with the estimate given in our obtained theorems.

Now we have

$$E\left[\int_0^T \frac{^nY_t}{n}dt\right] = \frac{1}{n}\int_0^T E[^nY_t]dt, \tag{5.2}$$

where $E[^nY_t]$ can be evaluated as follows. For the time increment h small enough, we have

$$\begin{aligned} E[^nY_{t+h}|^nY_t > 0] &= (^nY_t + 1)(nh\mu + o(h)) + (^nY_t - 1)(nh\mu + o(h)) \\ &\quad + ^nY_t(1 - nh\mu - nh\mu + o(h)) \end{aligned}$$

and

$$E[^nY_{t+h}|^nY_t = 0] = nh\mu + o(h),$$

where $o(h)$ stands for a term of order lower than h in the sense that $\lim_{h\to 0}\frac{o(h)}{h} = 0$.

This leads to

$$\begin{aligned}
& E[{}^nY_{t+h}|{}^nY_t] = {}^nY_t + nh\mu I\{{}^nY_t = 0\} + o(h) \\
\Rightarrow\ & E[{}^nY_{t+h}] = E[{}^nY_t] + nh\mu P\{{}^nY_t = 0\} + o(h) \\
\Rightarrow\ & \frac{E[{}^nY_{t+h}] - E[{}^nY_t]}{h} = n\mu P\{{}^nY_t = 0\} + \frac{o(h)}{h} \\
\Rightarrow\ & E[{}^nY_t] = \int_0^t n\mu P\{{}^nY_s = 0\}ds,
\end{aligned}$$

where the last step is a result of firstly taking $h \to 0$ and then integrating. Now continuing expression (5.2) we have

$$\begin{aligned}
E\left[\int_0^T \frac{{}^nY_t}{n}dt\right] &= \frac{1}{n}\int_0^T \int_0^t n\mu P\{{}^nY_s = 0\}ds\, dt \\
&= \mu \int_0^T \int_0^t e^{-2n\mu s}\{I_0(2n\mu s) + I_1(2n\mu s)\}ds\, dt, \qquad (5.3)
\end{aligned}$$

where the expression $P\{{}^nY_s = 0\} = e^{-2n\mu s}\{I_0(2n\mu s) + I_1(2n\mu s)\}$ with $I_0(2n\mu s)$ and $I_1(2n\mu s)$ standing for the modified Bessel functions of the first kind of order zero and one at $2n\mu s$ respectively, is in accordance with (Cantrell 1986, expression (1) there). Now let us set $\mu = T = 1$ (for simplicity), and recall that $\bar{T} = 1 = \gamma = N = L = 1$. Then the difference between our upper boundary estimate and the actual accuracy, given by (5.3), is plotted in the Figure 5.1; and the ratio of our estimate against the actual one is in Figure 5.2, both against the scaling parameter. From the figures, we see that our estimate is a rather rough one. In particular, Figure 5.2 suggests that the convergence happens around 30 times faster than estimated. In fact, when $n = 10,000$ the ratio of our estimate against the actual accuracy is about 32; and when $n = 100,000,000$ the ratio is also around 32. On the other hand, Figure 5.3 shows that the actual rate of convergence is of order $\frac{1}{\sqrt{n}}$, which is exactly our estimate. This says, given the uniformity in parameters as in Theorem 5.1, we do not have convergence faster than $\frac{1}{\sqrt{n}}$.

Finally, we emphasize two reasons for the settings of the example (assuming the queue to be empty initially and the unit traffic intensity under the tracking policy). The first reason is to have a simple expression for $P\{{}^nY_s = 0\}$, thus for $E\left[\int_0^T \frac{{}^nY_t}{n}dt\right]$, and hence for the actual rate of convergence. In fact, according to Cantrell (1986), for an arbitrary initial state n_0, denoting $\rho \stackrel{\triangle}{=} \frac{\lambda}{\mu}$ we have

$$\begin{aligned}
P\{{}^nY_s = 0\} &= e^{-(1+\rho)n\mu s}\Big\{\rho^{-n_0/2}I_{-n_0}(2\rho^{1/2}n\mu t) + \rho^{(-n_0-1)/2}I_{n_0+1}(2\rho^{1/2}n\mu s) \\
&\quad + (1-\rho)\rho^n \sum_{l=n_0+2}^{\infty} \rho^{-l/2}I_l(2\rho^{1/2}n\mu t)\Big\},
\end{aligned}$$

which is difficult for evaluation. Other alternative formulae, though available, see

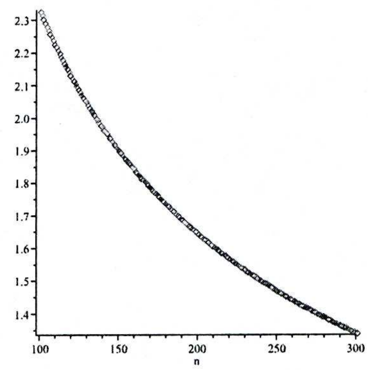

Figure 5.1: The difference between the actual and estimated accuracy for the example M/M/1 queue. The vertical axis stands for the difference, and the horizontal axis stands for the scaling parameter, natural n ranging from 100 to 300.

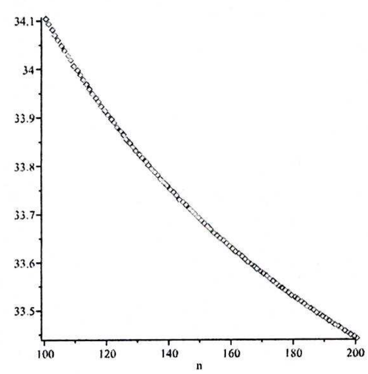

Figure 5.2: The ratio of the estimated against the actual accuracy for the example M/M/1 queue. The vertical axis stands for the ratio, and the horizontal axis stands for the scaling parameter, natural n ranging from 100 to 200.

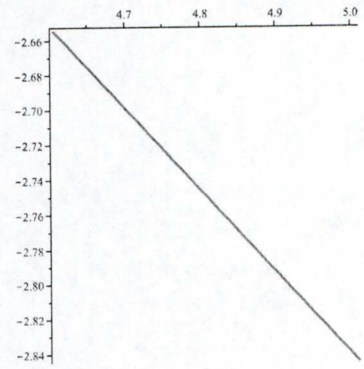

Figure 5.3: In the log scale: the actual accuracy for the example M/M/1 queue. The vertical axis stands for *log* of the actual accuracy, and the horizontal axis stands for $log(n)$, with natural n ranging from 100 to 150. Note that the curve is very close to a straight line with the slope of about $-1/2$.

(Sharma and Tarabia 2000, Sec.1) for example, are also rather complicated. In addition, when there are at least two subintervals, one deals with a time-dependent M/M/1 queue, for which the transient probability is provided in Zhang and Coyle (1991), as a solution to some Volterra-type integral equation, making it also very difficult to evaluate. After all, even in our simplified setting, the obtained expression for $|E\left[\int_0^T \frac{^nY_t}{n} dt\right]|$ is still not of a very simple form. However, at least, it is easy to compute with numerical values.

The second reason is that the simple setting itself is interesting and typical. Consider a bandwidth-sharing network with one resource and one flow. This is an $M/M/1$ queue, so that as before we feel free to use the languages of queueing theory. Recall the notations introduced in Section 5.2, and suppose the following: $Y_0 > 0$, $T = \frac{2Y_0}{\lambda}$ and the service rate can be controlled to take values from the interval $[0, 2\lambda]$. Then obviously the optimal fluid control is given by $^1\mu = 2\lambda$ so that the fluid model decreases at the fastest rate to 0 up to the time $\frac{T}{2}$ and $^2\mu = \lambda$ so that it stays at 0 from $\frac{T}{2}$ to T. Therefore, in particular, on $[\frac{T}{2}, T]$, we have exactly the example considered in the beginning of this section. On the other hand, applying the tracking policy, as we increase the scaling parameter, because the trajectory converges, see Chen and Mandelbaum (1991, 1994); Chen (1996); Dai (1995); Mandelbaum and Pats (1995) for example, at $\frac{T}{2}$ we are likely to end up with the stochastic model starting with the initial state close to zero as well as the unit traffic intensity, which, as we have seen in the above example, will result in the rate of convergence $O(\frac{1}{\sqrt{n}})$. This also explains one piece of intuition for avoiding fluid model to reach zero in the preabsorbing case we considered in Subsection 5.2.2.

5.4 Conclusion

To sum up, in this chapter, we have studied the efficiency of the tracking policy for the bandwidth-sharing network. While it was known for some other networks in the literature that the tracking policy is AFO, its efficiency, to our best knowledge, has not been studied intensively. Indeed, Gajrat et al (2003) is the only work we know that contains a short discussion on that topic. We have shown in terms of explicit upper boundary estimate of the rate of convergence for performance functionals that the tracking policy could be efficient or less depending on the parameters of the underlying fluid model. In particular, our result is in favor of tracking policy over the feedback type translations at least for short enough time horizons for its good efficiency as well as for the less information required. The current chapter contributes new insights about the accuracy of fluid approximations. It appears that the existing knowledge on the accuracy of fluid approximations is very scarce and needs significant development. We hope that the present chapter becomes an important step in this development.

5.5 Proof of main statements

Proof of Lemma 5.1. For the linear program (5.1), indeed, integrating by parts, we have the objective function in the following form:

$$\begin{aligned}
\int_0^T e'y(t)dt &= \int_0^T e'dt\, y(T) - \int_0^T \left(\int_0^t e'ds\right)(-Mu(t)+\Lambda)dt \\
&= \int_0^T e'dt(y(0) + \int_0^T (-Mu(t)+\Lambda)dt) \\
&\quad - \int_0^T \left(\int_0^t e'ds\right)(-Mu(t)+\Lambda)dt \\
&\quad (\text{write } \vec{t} \triangleq \textstyle\int_0^t e'ds = (t,\ldots,t)) \\
&= \vec{T}y(0) - \int_0^T \vec{T}Mu(s)ds + \vec{T}\Lambda T + \int_0^T \vec{t}Mu(t)dt - \int_0^T \vec{t}\Lambda dt.
\end{aligned}$$

Ignoring the uncontrolled terms, it is equivalent to considering the objective function

$$-\int_0^T \vec{T}Mu(s)ds + \int_0^T \vec{s}Mu(s)ds = \int_0^T (\vec{s}-\vec{T})Mu(s)ds.$$

Now, we note, $(\vec{s}-\vec{T})M$, being linear, is piecewise analytic on $[0,T]$, and all the other conditions required in (Pullan 1995, Thm.3.3) are satisfied, implying that the optimal control is a function $u(t)^*$ constant on intervals $[T_i, T_{i+1})$, where $i = 0, 1, \ldots, N-1 < \infty$ with $T_0 \triangleq 0$ and $T_N \triangleq T$. ■

Proof of Theorem 5.1

We observe that $\int_0^T e'\frac{^nY_t}{n}dt = \sum_{l=1}^{L}\int_0^T \frac{^nY_t^l}{n}dt$, meaning that under the fixed fluid optimal policy, we could effectively deal with only one "queue", and apply the same approach to the others [1]. Therefore, from now on we shall focus on files of type l, and consequently we have two one-dimensional processes $\{y_l(t), 0 \leq t \leq T\}$ and $\{^nY_t^l, 0 \leq t \leq T\}$, where we let $^nY_t^l$ and $y_l(t)$ indicate the lth element of ^nY_t and $y(t)$. For a real function $x(t)$, we denote $||x(t)||_{[t_1,t_2]} \triangleq \sup_{t\in[t_1,t_2]} x(t)$ and $||x(t)||_T \triangleq \sup_{t\in[0,T]} x(t)$, thus the uniform norm.

Lemma 5.2 *Consider a Poisson process with the counter* ^nA_t *and intensity* $n\lambda$. *Then* $P\{||\frac{^nA_t}{n} - \lambda t||_T \geq \varepsilon\} \leq \frac{\lambda T}{n\varepsilon^2}$.

Proof. Clearly, $\forall T > 0$, $M(t) \triangleq {}^nA_t - n\lambda t$ gives a martingale with right continuous trajectories (with the natural filtration) and the the index interval $[0,T]$. By the well-known Doob's L^p-inequality (Revuz and Yor 1999, Chap.2, Thm.1.7), $P\{||\frac{M(t)}{n}||_T \geq \varepsilon\} = P\{\sup_{0\leq t\leq T}|M(t)| \geq n\varepsilon\} \leq \frac{E[M(T)^2]}{(n\varepsilon)^2} = \frac{\lambda T}{n\varepsilon^2}$, as required. ■

Lemma 5.3 *Consider files of type l, thus the lth "queue" on the first subinterval. Recall the denotations introduced already in Section 5.2:* $^i\mu_l$ *is the service rate of the lth queue on the ith subinterval,* γ *is the maximum with respect to i and l of all* λ_l *and* $^i\mu_l$; *and* $\bar{T}$ *is the maximum length of all the N subintervals. Then* $P\{||\frac{^nY_t^l}{n} - y_l(t)||_{T_1} \geq \varepsilon\} \leq D_1 \triangleq \frac{72\bar{T}\gamma}{n\varepsilon^2}$.

Proof. Without generating confusion, we omit the index for the underlying "queue". Also without loss of generality (see also Remark 5.1 below), we assume $^1\mu > 0$. Define $^n\tilde{X}_t = {}^nY_0 + {}^nA_t - {}^nS_t$, where ^nS_t is a Poisson process with intensity $n\,^1\mu$, and ^nA_t is as in Lemma 5.2. Of course they are defined on the same probability space and independent. Define also $^n\hat{X}_t = \int_0^t I\{^nY_{s-} = 0\}d\,^nS_s$. Then it follows that $^n\hat{X}_t$ is nondecreasing and such that $\int_0^\infty I\{^nY_t > 0\}d\,^n\hat{X}_t = 0$. In this way we can write $^nY_t = {}^n\tilde{X}_t + {}^n\hat{X}_t$ in the form of the one-dimensional Skorokhod problem, where $^n\tilde{X}_t$ is the free process and $^n\hat{X}_t$ is the unused capacity process. This representation is also adopted in Mandelbaum and Pats (1995), and some other similar ones are in for example Pang and Day (2007); Robert (2003). The solution to this one-dimensional Skorokhod problem is well known, see (Mandelbaum and Pats 1995, Appendix A), so that we have $^nY_t = \varphi(^n\tilde{X}_t) = {}^n\tilde{X}_t + \sup_{0\leq s\leq t}[-(^n\tilde{X}_s \wedge 0)]$, where "$\wedge$" takes the minimum. Hence φ is a Lipschitz mapping with Lipschitz constant 2 on $D[0,\infty)$ equipped with the uniform norm. Here, $D[0,\infty)$ stands for the space of right continuous left limit functions on $[0,\infty)$. Indeed, for x_s,

[1] It is our standing assumption here that once a policy Φ is fixed, one flow is independent of another.

$y_s \in D[0,\infty)$,

$$\begin{aligned} ||\varphi(x_t) - \varphi(y_t)||_T &= ||x_t + \sup_{0\le s\le t}[-(x_s \bigwedge 0)] - y_t - \sup_{0\le s\le t}[-(y_s \bigwedge 0)]||_T \\ &\le ||x_t - y_t||_T + ||\sup_{0\le s\le t}[-(x_s \bigwedge 0)] - \sup_{0\le s\le t}[-(y_s \bigwedge 0)]||_T, \end{aligned}$$

where the second term in the last expression is clearly not bigger than $||x_t - y_t||_T$. Note also that the mapping φ is homogeneous in the sense that for any $x_s \in D[0,\infty)$, $\varphi(nx_s) = n\varphi(x_s)$. Define $\tilde{x}(t) = y(0) + (\lambda - {}^1\mu)t$, the deterministic analogue of ${}^n\tilde{X}_t$, and one can write down the Skorokhod problem for the fluid model as well.

Now we have

$$\begin{aligned} P\{||\frac{{}^nY_t}{n} - y(t)||_{T_1} \ge \varepsilon\} &= P\{||\varphi(\frac{{}^n\tilde{X}_t}{n}) - \varphi(\tilde{x}(t))||_{T_1} \ge \varepsilon\} \\ &\le P\{2||\frac{{}^n\tilde{X}_t}{n} - \tilde{x}(t)||_{T_1} \ge \varepsilon\} = P\{||\frac{{}^n\tilde{X}_t}{n} - \tilde{x}(t)||_{T_1} \ge \frac{\varepsilon}{2}\} \\ &\le P\{||\frac{{}^nY_0}{n} - y_0||_{T_1} \ge \frac{\varepsilon}{6}\} + P\{||\frac{{}^nA_t}{n} - \lambda t||_{T_1} \ge \frac{\varepsilon}{6}\} \\ &\quad + P\{||\frac{{}^nS_t}{n} - {}^1\mu t||_{T_1} \ge \frac{\varepsilon}{6}\} \qquad (5.4) \\ &\le \frac{36T\lambda}{n\varepsilon^2} + \frac{36T\,{}^1\mu}{n\varepsilon^2} \quad \text{(See Lemma 5.2)}, \end{aligned}$$

where the first equality and inequality are due to the homogeneity and Lipschitz continuity of φ. Now take γ to be the maximum of all λ and ${}^i\mu$ for all the L "queues", and $\bar{T}$ the maximum length of all the subintervals. Then we have $P\{||\frac{{}^nY_t}{n} - y(t)||_{T_1} \ge \varepsilon\} \le \frac{72\bar{T}\gamma}{n\varepsilon^2}$, as required. ■

Remark 5.1 *In the proof of Lemma 5.3, it was assumed that* ${}^1\mu_l > 0$. *Even if it fails to hold true, the same result still holds as then the third term in expression (5.4) will be absent.*

Corollary 5.1 *Recall that the optimal control is piece-wise constant with N subintervals. Then for any fixed $i = 1,\ldots,N$, on the ith subinterval, the following holds for files of any type l:* $P\{||\frac{{}^nY_t^l}{n} - y_l(t)||_{[T_{i-1},T_i]} \ge \varepsilon\} \le iD_1$.

Proof. Arguing similarly as in the proof of Lemma 5.3, we have that for files of type l on the second subinterval $P\{||\frac{{}^nY_t^l}{n} - y_l(t)||_{[T_1,T_2]} \ge \varepsilon\} \le D_1 + D_1 = 2D_1$, where the extra D_1 comes from the first term in the expression (5.4). D_1 is accumulated as we consider future intervals. The uniformity in l follows from the universal maximality of γ and $\bar{T}$, defined in the statement of Lemma 5.3. ■

Proof of Theorem 5.1. According to Corollary 5.1, we have

$$\begin{aligned} P\{\|\frac{{}^nY^l_t}{n} - y_l(t)\|_T \geq \varepsilon\} &\leq \sum_{i=1}^{N} P\{\|\frac{{}^nY^l_t}{n} - y_l(t)\|_{[T_{i-1},T_i]} \geq \varepsilon\} \\ &= \frac{D_1 N(1+N)}{2}. \end{aligned} \tag{5.5}$$

Note that the above estimate is uniform in l. Now

$$\begin{aligned} |{}^n\hat{W} - \hat{w}| &= |E_{{}^nY_0}\int_0^T \sum_{l=1}^{L} \frac{{}^nY^l_t}{n} dt - \int_0^T \sum_{l=1}^{L} y_l(t) dt| \leq \sum_{l=1}^{L} E_{{}^nY_0}\int_0^T |\frac{{}^nY^l_t}{n} - y_l(t)| dt \\ &= \sum_{l=1}^{L}\int_0^T E_{{}^nY_0}|\frac{{}^nY^l_t}{n} - y_l(t)| dt = \sum_{l=1}^{L}\int_0^T\int_0^\infty P_{{}^nY_0}\{|\frac{{}^nY^l_t}{n} - y_l(t)| \geq \varepsilon\} d\varepsilon\, dt \\ &\leq \sum_{l=1}^{L}\int_0^T \left\{ \int_0^{\frac{6\sqrt{N(N+1)\gamma\bar{T}}}{\sqrt{n}}} d\varepsilon + \int_{\frac{6\sqrt{N(N+1)\gamma\bar{T}}}{\sqrt{n}}}^\infty \frac{36N(N+1)\gamma\bar{T}}{n\varepsilon^2} d\varepsilon \right\} dt \\ &\quad \text{(See expression (5.5))} \\ &= \frac{TL}{\sqrt{n}}\left\{12\sqrt{N(N+1)\gamma\bar{T}}\right\}, \end{aligned}$$

where the second equality, the interchange of integrals, is a result of Tonelli's theorem, and the third equality is simply the telescope formula, as required. Note in the second last line, we put the lower limit and upper limit of the two integrals $\frac{C}{\sqrt{n}}$ in order to get the fastest possible convergence of order $\frac{1}{\sqrt{n}}$, and $C = 6\sqrt{N(N+1)\gamma\bar{T}}$ for the same reason. ■

Proof of Theorem 5.2

Let us introduce some additional notations. We shall denote $\hat{w}^l$ and ${}^n\hat{W}^l$ the lth summand of $\hat{w}$ and ${}^n\hat{W}$. That is, $\hat{w}^l = \int_0^T y_l(t)dt$ and ${}^n\hat{W}^l = E_{Y_0^l}[\int_0^T \frac{{}^nY^l_t}{n} dt]$. Again, the indications of control policies are omitted, as they are fixed. Define for $t < T$, $w^l(x,t) = \int_t^T \tilde{y}_l(s)ds$, $\tilde{y}_l(t) = x$, while $w^l(x,t) \triangleq 0$ for $t \geq T$. Here $\tilde{y}_l$ is different from y_l only because the former ignores the nonnegativity constraint, or in other words, can be negative. We shall modify w^l such that it becomes null for all its state argument bigger than Y. That is, we define $v^l(x,t) \triangleq w^l(x,t)I\{x \leq Y\}$.

Lemma 5.4 *There exist an n-independent positive constant C_l and positive function ζ_l such that for any $\varepsilon > 0$, $P_{{}^nY_0^l}\{\|\frac{{}^nY^l_t}{n} - y_l(t)\|_T \geq \varepsilon\} \leq C_l e^{-\zeta_l(\varepsilon)n}$.*

Proof. According to Lemma 5.1, the process ${}^nY^l_t$ is stationary on intervals $[0,T_1)$, $\ldots$, $[T_{N-1},T]$. Moreover, under the fixed policy, on each interval, the process has arrival and potential service processes both renewal. Consider now the first subinterval. Then by (Chen 1996, Thm.3.1) we have an n-independent positive constant ${}^1C_l$ and positive function ${}^{1}\zeta_l$ such that $P\{||\frac{{}^{n}Y_t^l}{n} - y_l(t)||_{T_1} \geq \varepsilon\} \leq {}^{1}C_l e^{-{}^{1}\zeta_l(\varepsilon)n}$. Here it is convenient to omit the subscript and write simply P. In particular, we have $P\{|\frac{{}^{n}Y_{T_1}^l}{n} - y_l(T_1)| \geq \varepsilon\} \leq {}^{1}C_l e^{-{}^{1}\zeta_l(\varepsilon)n}$.

Keeping in mind the last inequality, if we take ${}^{n}Y_{T_1}$ as the initial state of the process on the next interval, then the above reasoning can again be applied, and so on. Eventually, there will be N positive ${}^{1}C_l$-like constants and ${}^{1}\zeta_l(\varepsilon)$-like functions.

Finally, we have

$$
\begin{aligned}
P\{||\frac{{}^{n}Y_l}{n} - y(t)||_T \geq \varepsilon\} &\leq \sum_{i=1}^{N} P\{||\frac{{}^{n}Y_l}{n} - y(t)||_{[T_{i-1},T_i]} \geq \varepsilon\} \leq \sum_{i=1}^{N} {}^{i}C_l e^{-{}^{i}\zeta_l(\varepsilon)n} \\
&\leq C_l e^{-\zeta_l(\varepsilon)n},
\end{aligned}
$$

where $\zeta_l(\varepsilon) \triangleq \min_{i=1,\dots,N} {}^{i}\zeta_l(\varepsilon)$ and $C_l \triangleq \sum_{i=1}^{N} {}^{i}C_l$. ■

Remark 5.2 *In the above lemma, the exact expressions for the constant C_l and the function Φ_l can be obtained by examining the proof of (Chen 1996, Thm.2.2, 2.3, 3.1). However, the obtained expressions are much more complicated than such obtained in (5.5), and are difficult to be used in the proof of Theorem 5.1. On the other hand, the exponentially fast decrease is needed in the proof of Theorem 5.2 below.*

Lemma 5.5 *For all l, the following assertions hold on the interval $[0,T)$:*
(a) w^l solves the following equation of dynamic programming (DP) type,

$$
-\frac{\partial w^l}{\partial t}(\tilde{y}_l,t) = \tilde{y}_l + \frac{\partial w^l}{\partial \tilde{y}_l}(\tilde{y}_l,t)\frac{d\tilde{y}_l}{dt}, \tag{5.6}
$$

with the boundary condition $w^l(\tilde{y}_l,T) = 0$. Here we take $\frac{\partial w^l(\tilde{y}_l,t)}{\partial t}$ at T_i, $i = 0\dots,N-1$ as the right derivative. In particular, $\hat{w}^l = w(Y_0^l,0)$.
(b) There is $\beta > 0$ so that $w^l(\tilde{y}_l,t)$, $\frac{\partial w^l}{\partial \tilde{y}_l}(\tilde{y}_l,t)$, $\frac{\partial w^l}{\partial t}(\tilde{y}_l,t)$, $\frac{\partial^2 w^l}{\partial \tilde{y}_l^2}(\tilde{y}_l,t)$ are all bounded by β on $[0,Y]\times[0,T]$, uniformly in l.
(c) $v^l(\tilde{y}_l,t)$ is bounded and has a bounded partial derivative $\frac{\partial v^l(\tilde{y}_l,t)}{\partial t}$. Here we take $\frac{\partial v^l(\tilde{y}_l,t)}{\partial t}$ at T_i, $i = 0\dots,N-1$ as the right derivative.

Proof. The exact expression for w^l can be obtained easily. On the other hand, there are finitely many points where $w^l(\tilde{y}_l,t)$ does not have the partial derivative with respect to t. Parts (a) and (b) of the lemma can then be directly verified using inductive arguments. Part (c) follows from Parts (a) and (b) and the definition of v^l. ■

Since Φ^* is non-stationary, the resulting process is non-stationary, and we shall apply the extended generator $\tilde{L}$, whose definition can be found in (Hernández-Lerma 1994, Chap.1). Here, we refer readers to the references for this definition because to introduce it, many extra denotations must be introduced for such as transition functions, semigroups, shifted operators and so on. We have the following lemma describing $\tilde{L}$.

Lemma 5.6 *$v^l(i,t)$ is in the domain of $\tilde{L}$, with the form of*

$$\tilde{L}v^l(i,t) = \frac{\partial v^l}{\partial t}(i,t) + \sum_{j \in S} q(j|i,t)v^l(j,t),$$

where $[q(j|i,t)]_{i,j=0,1,\ldots}$ gives the Q matrix of the corresponding (unique) continuous-time Markov chain corresponding to the lth queue.

Proof. One can simply examine the proof of (Hernández-Lerma 1994, Prop.14.4). ■

Proof of Theorem 5.2. In what follows, without generating confusion, apart from the index standing for control policy, we shall omit the index corresponding to the type of files. There is no confusion if we drop the index standing for which queue we are considering, and put $q(i+1|i,t) = \lambda(t) = \lambda \leq \bar{\lambda}$ and similarly $q(i-1|i,t) = \mu(t) = \Phi^*({}^nY_t,t)\mu \leq \bar{\mu}$. Here let us emphasize that λ and $\lambda(t)$ (μ and $\mu(t)$) are different, in case there is any ambiguity. Due to Lemma 5.6 and (Hernández-Lerma 1994, Lem.2.1), the following Dynkin's formula is valid:

$$0 \quad = \quad v({}^nY_0/n,0) + E_{{}^nY_0}[\int_0^T \tilde{L}v({}^nY_t/n,t)ds]. \tag{5.7}$$

Adding $E_{{}^nY_0}[\int_0^T {}^nY_t/n\,dt]$ to the both sides and some arrangements give

$${}^n\hat{W} - v({}^nY_0/n,0) = E_{{}^nY_0}[\int_0^T \frac{{}^nY_t}{n} + \tilde{L}v(\frac{{}^nY_t}{n},t)dt]. \tag{5.8}$$

This is justified by the fact that $E_{{}^nY_0}[\int_0^T {}^nY_t/n\,dt]$ is bounded by a function of ${}^nY_0 < \infty$. Recall that the scaled queue is stochastically dominated by a Poisson process with intensity $n\lambda$. In addition, we remind, $\hat{w} = w(Y_0,0) = v(Y_0,0)$.

Next we shall analyze (5.8) in three cases.

(i) Suppose the instantaneous state ${}^nY_t$ is positive and smaller than nY.

Writing out the right hand side term in (5.8), we have the integrand in the form

$$\begin{aligned}
&\frac{{}^nY_t}{n} + \frac{\partial v({}^nY_t/n,t)}{\partial t} + n\lambda v(\frac{{}^nY_t+1}{n},t) + n\mu(t)v(\frac{{}^nY_t-1}{n},t) - n(\lambda+\mu(t))v(\frac{{}^nY_t}{n},t) \\
= &\frac{{}^nY_t}{n} + \frac{\partial v({}^nY_t/n,t)}{\partial t} + n\lambda(v(\frac{{}^nY_t+1}{n},t) - v(\frac{{}^nY_t}{n},t)) \\
&+ n\mu(t)(v(\frac{{}^nY_t-1}{n},t) - v(\frac{{}^nY_t}{n},t)).
\end{aligned} \tag{5.9}$$

Recall the equation of DP type (5.6), rewrite $\frac{{}^nY_t}{n}$ accordingly, and recall the fact that $w({}^nY_t,t) = v({}^nY_t,t)$ for $0 < {}^nY_t < nY$ due to the definition of v. We find that the partial derivative with respect to t is cancelled out and the rest of the right hand side of (5.9) becomes

$$RHS \text{ of } (5.9) = n\lambda(v(\frac{^nY_t+1}{n},t) - v(\frac{^nY_t}{n},t))$$

$$+n\mu(t)(v(\frac{^nY_t-1}{n},t) - v(\frac{^nY_t}{n},t)) - (\lambda - \mu(t))\frac{\partial v}{\partial i}(^nY_t/n,t).$$

Since our aim is to estimate $|^n\hat{W} - \hat{w}|$, we shall take modulus inside the latter expression and then take it as the integrand. As a result, we have

$$\begin{aligned}
&|E_{^nY_0}[\int_0^T (\frac{^nY_t}{n} + \tilde{L}v(\frac{^nY_t}{n},t))I\{^nY_t \in (0,nY)\}dt]| \\
\leq\ & E_{^nY_0}\int_0^T \Big\{\bar{\lambda}|n(v(\frac{^nY_t+1}{n},t) - v(\frac{^nY_t}{n},t)) - \frac{\partial v}{\partial i}(\frac{^nY_t}{n},t)| \\
&+\bar{\mu}|\frac{\partial v}{\partial i}(\frac{^nY_t}{n},t) - n(v(\frac{^nY_t}{n},t) - v(\frac{^nY_t-1}{n},t))|I\{^nY_t \in (0,nY)\}\Big\}dt \\
\leq\ & \frac{(\bar{\lambda}+\bar{\mu})\beta T}{2n}, \qquad (5.10)
\end{aligned}$$

where the last inequality is attained by applying Taylor's theorem to $v(\frac{^nY_t}{n},t) - v(\frac{^nY_t-1}{n},t)$ and $v(\frac{^nY_t+1}{n},t) - v(\frac{^nY_t}{n},t)$, bounding the indicator by 1 and using Lemma 5.5.

(ii) Suppose the instantaneous state $^nY_t/n = 0$.

Then there will not be $\mu(t)$ in the integrand, as the state cannot be negative. That is, with the substitution of $^nY_t/n = 0$,

$$\begin{aligned}
&E_{^nY_0}[\int_0^T (\frac{^nY_t}{n} + \tilde{L}v(\frac{^nY_t}{n},t))I\{^nY_t = 0\}dt] \\
=\ & E_{^nY_0}[\int_0^T (\frac{\partial v(0,t)}{\partial t} + n\lambda v(\frac{1}{n},t) - n\lambda v(0,t))I\{^nY_t = 0\}dt] \\
\leq\ & \beta(1+2n\bar{\lambda})E_{^nY_0}\int_0^T I\{^nY_t = 0\}dt \\
=\ & \beta(1+2n\bar{\lambda})E_{^nY_0}[\text{ time spent by } {^nY_t} \text{ at 0 by } T] \\
=\ & \beta(1+2n\bar{\lambda})E_{^nY_0}[\text{ time spent by } \tfrac{^nY_t}{n} \text{ at 0 by } T] \\
=\ & \beta(1+2n\bar{\lambda})E_{^nY_0}[\text{ time spent by } \tfrac{^nY_t}{n} \text{ at 0 by } T | \tfrac{^nY_t}{n} \text{ visits 0 by } T] \\
& \times P_{^nY_0}\{\tfrac{^nY_t}{n} \text{ visits 0 by } T\} \\
\leq\ & \beta(1+2n\bar{\lambda})TP_{^nY_0}\{||\frac{^nY_t}{n} - y(t)||_T \geq \underline{y}\} \\
\leq\ & \beta(1+2n\bar{\lambda})TC_t e^{-\zeta_t(\underline{y})n}, \qquad (5.11)
\end{aligned}$$

where the last inequality follows from Lemma 5.4. Here, clearly we see the upper bound for $P_{^nY_0}\{||\frac{^nY_t}{n} - y(t)||_T \geq \underline{y}\}$ based on (5.5), which converges to 0 as fast as $\frac{1}{n}$ is not enough.

(iii) Suppose the instantaneous state $^nY_t/n = Y$.

Then

$$E_{^nY_0}[\int_0^T (\frac{^nY_t}{n} + \tilde{L}v(\frac{^nY_t}{n},t))I\{^nY_t = nY\}dt]$$

$$\begin{aligned}
&= E_{^nY_0}[\int_0^T (Y + \frac{\partial v(Y,t)}{\partial t} + n\mu(t)v(\frac{nY-1}{n},t) - n(\lambda+\mu(t))v(Y,t))I\{^nY_t = nY\}dt] \\
&\leq (Y+\beta+n\bar{\lambda}\beta+2n\bar{\mu}\beta)E_{^nY_0}\int_0^T I\{\frac{^nY_t}{n} = Y\}ds \\
&= (Y+\beta+n\bar{\lambda}\beta+2n\bar{\mu}\beta)E_{^nY_0}[\text{ time spent by } \tfrac{^nY_t}{n} \text{ at } Y \text{ by } T] \\
&= (Y+\beta+n\bar{\lambda}\beta+2n\bar{\mu}\beta)E_{^nY_0}[\text{ time spent by } \tfrac{^nY_t}{n} \text{ at } Y \text{ by } T | \tfrac{^nY_t}{n} \text{ visits } Y \text{ by } T] \\
&\quad \times P_{^nY_0}\{\tfrac{^nY_t}{n} \text{ visits } Y \text{ by } T\} \\
&\leq (Y+\beta+n\bar{\lambda}\beta+2n\bar{\mu}\beta)TP_{^nY_0}\{||\frac{^nY_t}{n} - y(t)||_T \geq Y - \bar{y}\} \\
&\leq (Y+\beta+n\bar{\lambda}\beta+2n\bar{\mu}\beta)TC_l e^{-\zeta_l(Y-\bar{y})n}, \qquad (5.12)
\end{aligned}$$

where the last inequality follows from Lemma 5.4.

Finally, adding expressions (5.9, 5.11, 5.12) up will result in

$$|^n\hat{W} - \hat{w}| \leq \frac{(\bar{\lambda}+\bar{\mu})\beta T}{2n} + \beta(1+2n\bar{\lambda})TC_l e^{-\zeta_l(\underline{y})n} + (Y+\beta+n\bar{\lambda}\beta+2n\bar{\mu}\beta)TC_l e^{-\zeta_l(\bar{y})n}.$$

Recall this estimate is for the queue of files of type l. Hence we shall denote $\eta^l(n) \triangleq \frac{(\bar{\lambda}+\bar{\mu})\beta T}{2n} + \beta(1+2n\bar{\lambda})TC_l e^{-\zeta_l(\underline{y})n} + (Y+\beta+n\bar{\lambda}\beta+2n\bar{\mu}\beta)TC_l e^{-\zeta_l(\bar{y})n}$. The reason to indicate the file type l here lies in the fact that the form of the function ζ_l depends on which queue is under consideration.

Repeating the above reasoning $L-1$ times will result in another $L-1$ η^l-like functions. Adding the differences together will result in

$$|^n\hat{W} - \hat{w}| \leq \sum_{l=1}^{L} \eta^l(n) = C \cdot O(\frac{1}{n}),$$

where C is some constant. ■

Chapter 6

Fluid model of an Internet router[1]

6.1 Introduction

Most traffic in the Internet is governed by TCP/IP (Allman et al (1999); Jacobson (1988)). TCP protocol tries to adjust the sending rate of a source to match the available bandwidth along the path. The current TCP New Reno uses MIMD congestion control during the initial Slow Start phase and AIMD congestion control during the principal Congestion Avoidance phase. In the AIMD congestion control scheme in the absence of congestion signals from the network, TCP enlarges the congestion window linearly in RTTs (Round Trip Times) and, upon the reception of a congestion signal, TCP reduces the congestion window by a multiplicative factor. In the MIMD congestion control scheme in the absence of congestion signals from the network, TCP enlarges the congestion window exponentially in RTTs.

A significant increase of link capacities has posed a challenge to the current TCP implementation. The current TCP New Reno version is not able to utilize efficiently high speed links (Floyd (2003)). To mitigate this problem, several new TCP versions (HS-TCP, FAST-TCP, Scalable TCP, H-TCP, CUBIC-TCP, BIC-TCP for example) have been proposed, see Floyd (2003); Jin et al (2004); Kelly (2003); Leith and Shorten (2004); Rhee and Xu (2005); Xu et al (2004). These algorithms have in common that in the absence of congestion, the sources enlarge the congestion window in a much more aggressive fashion than the standard TCP New Reno does. An extensive overview and comparison of different TCP versions for high capacity links is given in Li et al (2007). In the present chapter we analyze the MIMD congestion control which is a base for Scalable TCP (Kelly (2003)).

[1]The results presented in this chapter are based on Zhang et al (2010).

On the other hand, most of the routers in the Internet are of Drop Tail type. In basic Drop Tail routers, apart from the router capacity, the buffer size is the only parameter to be tuned. In fact, the buffer size is one of the few parameters of the TCP/IP network that can be managed by network operators. This makes the choice of the router buffer size very important in the TCP/IP network design. This choice has recently received considerable attention, see Appenzeller et al (2004); Avrachenkov et al (2005, 2010, 2002); Ayesta et al (2008); Enachescu et al (2005); Gorinsky et al (2005); Prasad et al (2007); Raina et al (2005); Raina and Wischik (2005); Stanojević et al (2006); Villamizar and Song (1994); Vu-Brugier et al (2007); Wischik and McKeown (2005). (This is far from an exhaustive list of relevant references.) However, most of these works study only the AIMD congestion control algorithm.

In this chapter, we study the interaction of MIMD congestion control algorithms with Drop Tail buffers. We consider the problem in the framework of deterministic hybrid models, which describe systems with both discrete and continuous behavior. Recently, hybrid models have been successfully applied to the modeling of communication networks (Avrachenkov et al (2005, 2010); Bohacek et al (2003); Hespanha et al (2001)). The model in the present chapter is a significant extension of the models in Ayesta et al (2008)[2]. In particular, in Ayesta et al (2008), the RTT is regarded ignorably small, so that there is no delay between sending data out and receiving the corresponding acknowledgements. This means that as soon as the buffer is filled full, there will be an instantaneous multiplicative reduction (without any delay) on the sending rate. In comparison, in the current chapter, as will be seen in Section 6.2, we take accurately into account the time-varying nature of the RTT, resulting in a time-varying delay between sending out data and receiving corresponding acknowledgements. The present more accurate model allows us to provide conditions for the absence of multiple subsequent reductions of the congestion window and estimate more accurately the minimal buffer size for the full link utilization. Furthermore, we recommend the use of the Delayed Ack mechanism (Allman et al (1999)) and the reduction of the window growth parameter in order to avoid the undesirable regime of subsequent window reductions. Additionally to the analytical expression for the minimal buffer size for the full link utilization, we construct the Pareto set to achieve the trade off between the high link utilization and small queueing delays. In particular, our results suggest that in order to achieve high utilization, one can size the buffer much smaller than the bandwidth-delay product. Our analytical results are confirmed by NS simulations.

The rest of this chapter is organized as follows: we describe our mathematical model in Section 6.2. In Section 6.3 we state the result regarding system trajectories, while Section 6.4 is about the results on optimal buffer sizing. The results are verified in Section 6.5 by means of simulations. We finish this chapter with a conclusion. The

[2] We studied the model in Ayesta et al (2008) at an early stage of this research, and that study is included in Appendices for completeness and convenient comparisons.

proofs of the main statements are collected in Section 6.7.

6.2 Mathematical model

Consider a long-lived MIMD TCP connection that sends data through a bottleneck router. Denote by $w(t)$ the instantaneous congestion window of the TCP connection at time $t \in [0, \infty)$. Let $x(t)$ be the amount of data in the bottleneck queue at time t, $B > 0$ be the size of the Drop Tail buffer, and μ be the capacity of the bottleneck router.

If $x(t) < B$, the evolution of $w(t)$ is given by differential equation

$$\frac{dw}{dt} = \frac{mw}{T + x(t)/\mu}. \tag{6.1}$$

Here T is the two way propagation delay and m being a constant, is some fixed multiplicative factor. Note that $T + x(t)/\mu$ corresponds to the RTT at time moment t.

The sending rate of the window based congestion control is given by

$$\lambda(t) = \frac{w(t)}{T + x(t)/\mu}. \tag{6.2}$$

We emphasize that the time parameter t corresponds to the local time observed at the router.

When x reaches B at time t^*, i.e. $x(t^*) = B$, the buffer starts to overflow. The overflow of the buffer will be noticed by the sender only after the time delay $\delta = T + B/\mu$. Upon the reception of the congestion signal at time $t^* + \delta$, the congestion window is reduced according to

$$w(t^* + \delta + 0) = \beta^k w(t^* + \delta - 0), \tag{6.3}$$

where $k \stackrel{\triangle}{=} \min\{i = 1, 2, \cdots : \beta^i w(t^* + \delta - 0) < \mu(T + \frac{B}{\mu})\}$. Usually, $k = 1$, but sometimes it is necessary to send several congestion signals in order to reduce the sending rate below the transmission capacity of the bottleneck router.

Therefore, between the instantaneous jumps of the congestion window w, we have the dynamical system

$$\dot{x} = \begin{cases} \lambda(t) - \mu, & \text{if } 0 < x(t) < B, \text{ or } x(t) = 0 \text{ and } \lambda(t) \geq \mu, \\ & \text{or } x(t) = B \text{ and } \lambda(t) \leq \mu; \\ 0 & \text{otherwise,} \end{cases} \tag{6.4}$$

where $\lambda(t)$ is given by (6.2).

Let us discuss particular parameter settings. Currently, the MIMD congestion control mechanism is used in:

(a) Slow Start regime (Allman et al (1999)) in the standard TCP New Reno;

(b) Scalable TCP (Kelly (2003)) for high speed links.

In the Slow Start regime we have $\beta = 0.5$. The value of m depends on whether the Delayed Ack mechanism (Allman et al (1999)) is enabled or not. If the Delayed Ack mechanism is enabled, $m = 0.5$, and if it is not enabled, $m = 1$.

In Scalable TCP we have $\beta = 0.875$ and $m = 0.01$.

We would like to recall that a similar hybrid model can be used to study the AIMD congestion control (Avrachenkov et al (2010); Bohacek et al (2003); Hespanha et al (2001)). One only needs to change equation (6.1) to the following equation

$$\frac{dw}{dt} = \frac{M}{T + x(t)/\mu}.$$

The AIMD congestion control is used in the principal Congestion Avoidance regime of TCP New Reno. In this case, we have $\beta = 0.5$, and M is equal to the half packet size if the Delayed Ack mechanism is enabled, and otherwise M is equal to the packet size.

6.3 Convergence to limiting cycles

Let us first begin with some definitions.

Definition 6.1 *A cycle is defined as the trajectory starting with the initial state $w(0) = w_0 = W_0 \in [\beta(\mu T + B), \mu T + B)$, $x(0) = x_0 = B$ at $t = 0$, and reaching the same point for the first time at some moment T_{cycle}, called the duration of the cycle. Note that $T_{cycle} > \delta = T + B/\mu$ because $W_0 < B + \mu T$. A cycle with $x(t)$ staying at zero for a positive time interval is called clipped. Otherwise it is unclipped. In particular, a cycle with $x(t)$ staying at 0 at a single time moment is called critical, and it is referred to as an unclipped cycle. In addition, a cycle, possibly with more than one instantaneous jump though (i.e. $k > 1$ in (6.3)), is called simple, if it has only one loop (one convex time interval containing no jumps (6.3)). Otherwise, it is called complicated.*

The case of a simple cycle with $k = 1$ is most interesting because in this case we avoid multiple subsequent packet losses. Such a cycle will be called a $1-$cycle or a cycle of order one. In the general case, a simple cycle is called $k-$cycle (a cycle of order k).

Let

$$B^* = \mu T \frac{1 - e^{m+1}\beta^{\frac{m+1}{m}} - (m+1)e^{m}\left(1 - e\beta^{\frac{1}{m}}\right)\beta e^{ms_1}}{(m+1)e^{m}\left(1 - e\beta^{\frac{1}{m}}\right)\beta e^{ms_1}}, \tag{6.5}$$

where

$$s_1 = \frac{1}{m+1} \ln \left(\frac{1 - e^m \beta}{\beta e^m m \left(1 - e\beta^{\frac{1}{m}}\right)} \right). \tag{6.6}$$

Theorem 6.1 *(a) For an arbitrary $B > B^*$, the system trajectory converges to the limiting unclipped 1-cycle from an arbitrary initial state iff $\beta < \bar{\beta}$, where $\bar{\beta}$ is the single solution to*

$$(m+1)e^m \bar{\beta}(1 - e\bar{\beta}^{\frac{2}{m}}) + e^{m+1}\bar{\beta}^{2(1+\frac{1}{m})} = 1 \tag{6.7}$$

in the interval $(0, e^{-m})$.
(b) Suppose $B = B^*$. Then the limiting cycle is of order one and critical iff $\beta < e^{-m}$.
(c) Suppose $B < B^*$. Then the limiting cycle is of order one and clipped iff $\beta < e^{-m}$.

In cases (b) and (c), the system trajectory also converges to the limiting cycle from an arbitrary initial state.

A simple $1-$cycle (clipped or unclipped) exists iff $\beta e^m < 1$.

Remark 6.1 *According to the proofs given in Section 6.7, in case (a), condition $\beta < \bar{\beta}$ can be relaxed to $\beta < e^{-m}$ sacrificing the convergence from an arbitrary initial state. Specifically, if $\beta < e^{-m}$, $B > B^*$, and $w_0 \in [\beta e^m (B + \mu T), B + \mu T)$, then the system trajectory converges to the limiting cycle, which is of order one and unclipped.*

Inequality $\beta < \bar{\beta}$ is a sufficient condition for the convergence from an arbitrary initial state in all three cases of Theorem 6.1.

Suppose $\bar{\beta} \leq \beta < e^{-m}$. According to the proof of Theorem 6.1, in case $B > B^*$ a trajectory does not converge to the limiting unclipped $1-$cycle iff after each series of jumps $w(t^* + \delta + 0) < \beta e^m (B + \mu T)$. In this situation, double jumps always happen, so that one can use the developed theory with β being replaced with β^2. As a result, one can face only the convergence to a simple $2-$cycle which can be unclipped, critical or clipped. Complicated cycles never appear.

In particular, the above theorem implies that the buffer size B^* is the minimal buffer size for the full link utilization. The following asymptotics holds for small values of m

$$B^*(m) = \mu T \frac{(1 - \beta + m\ln(m))}{\beta} + o(m\ln(m)), \tag{6.8}$$

where $o(m\ln(m))$ vanishes faster than $m\ln(m)$. The asymptotics (6.8) can be verified by the application of the L'Hôpital's rule (See Lemma 6.3 in Section 6.7). The asymptotics (6.8) together with the exact expression (6.5) can be considered as an improvement of the results presented in Ayesta et al (2008); Khoury and Altman (2004). In particular, for Scalable TCP the above asymptotics gives $B^* \approx 0.09\mu T$. Thus, a single Scalable TCP connection requires about 10 times less buffer space than a standard TCP New Reno connection, which requires up to μT buffer space (Villamizar and Song (1994)).

Note that in the Slow Start phase of TCP New Reno without the Delayed Ack mechanism (Allman et al (1999)) the condition $\beta < e^{-m}$ is violated and it is possible to have subsequent window reductions. However, if the Delayed Ack mechanism is enabled, the value of m reduces from 1 to 0.5 and condition $\beta < e^{-m}$ is satisfied. We note that if $m = 0.5$, condition $\beta < \bar{\beta}$ is not satisfied since in this case $\bar{\beta} = 0.43$. However, even if condition $\beta < \bar{\beta}$ is violated, the system trajectory can still converge to 1-cycle from some initial conditions. To avoid for sure the undesired regime of multiple window reductions, one can reduce the value of m to 0.4 in the Slow Start regime.

In the case of Scalable TCP, the inequality $\beta < \bar{\beta}$ is valid as $\bar{\beta} \approx 0.98$, and the regime of multiple window reductions is not realized in any network conditions and configurations.

6.4 Pareto set for the buffer sizing

Let us study what effect has the choice of the buffer size on the performance of TCP with MIMD congestion control. In particular, we are interested in the optimal buffer sizing. We have two criteria here, namely the average throughput, defined by

$$\bar{g} = \lim_{t \to \infty} \frac{1}{t} \int_0^t g(t)dt,$$

where

$$g(t) = \begin{cases} \lambda(t) & \text{if } x(t) < B \\ \mu & \text{if } x(t) = B, \end{cases}$$

and the average amount of data in the buffer, defined by

$$\bar{x} = \lim_{t \to \infty} \frac{1}{t} \int_0^t x(t)dt.$$

More precisely, one is interested in maximizing $\bar{g}$ and minimizing $\bar{x}$. Clearly those two objectives are contradictory. This is a typical situation in multi-criteria optimization. A standard approach is to optimize one criterion under constraints on the other one. And the solution providing the optimality gives a point in the Pareto set. As is known, see e.g. Piunovskiy (1997), it can be obtained by solving the optimization problem

$$\max_B \left\{ \lim_{t \to \infty} \frac{1}{t} \int_0^t (c_1 g(t) - c_2 x(t))dt \right\}$$

with $(c_1, c_2) \in \mathbf{R}^2_+$. Different values of $c_1 > 0$ and $c_2 > 0$ lead to the complete Pareto set which must be closed. Based on the Pareto set, one can make the decision on the parity between the two objectives. Mathematical description of partial orders and connected Pareto sets can be found in Dorini et al (2007).

We study the Pareto optimality in the framework of the simple clipped (or critical) $1-$cycle, i.e. we assume that $\beta < e^{-m}$ and $B \leq B^*$. The formulae for $\bar{g}$ and $\bar{x}$ can be written as

$$\bar{g} = \frac{1}{T_{cycle}} \int_0^{T_{cycle}} g(t)dt,$$

and

$$\bar{x} = \frac{1}{T_{cycle}} \int_0^{T_{cycle}} x(t)dt,$$

where T_{cycle} is the duration of the cycle. The following propositions provide expressions for the average sending rate, throughput, and amount of data in the buffer. In particular, the expressions allow us to plot the Pareto set parameterized by the buffer size.

Firstly, consider the case $B \geq B^*$ and suppose the limiting $1-$cycle is realized (see Theorem 6.1(a)). Then the duration of that cycle equals

$$T_{cycle} = \frac{B+\mu T}{\mu m} \left\{ m + \frac{(1-e\beta^{\frac{1}{m}})(m+1)(1-e^m\beta)}{1-e^{m+1}\beta^{1+\frac{1}{m}}} \right\}, \tag{6.9}$$

and the following proposition holds.

Proposition 6.1 *The average sending rate is given by*

$$\bar{\lambda} = \frac{(1-\beta)e^m(1-e\beta^{1/m})(m+1)\mu}{\left[m(1-e^{m+1}\beta^{1+1/m}) + (1-e\beta^{1/m})(m+1)(1-e^m\beta)\right]}, \tag{6.10}$$

the average throughput is given by

$$\bar{g} = \mu,$$

and the average amount of data in the buffer is given by

$$\bar{x} = \frac{1}{T_{cycle}} \left\{ TB \int_0^S y(s)ds + \frac{B^2}{\mu} \int_0^S y^2(s)ds + BT + \frac{B^2}{\mu} \right\}, \tag{6.11}$$

where $S = \frac{1}{m} \ln \frac{1}{\beta} - 1$, *and*

$$y(s) = \frac{\beta(1+\frac{\mu T}{B})e^m(1-e\beta^{1/m})}{1-e^{m+1}\beta^{1+1/m}}(e^{ms}-e^{-s}) - \frac{\mu T}{B} + e^{-s}(1+\frac{\mu T}{B}).$$

Secondly, consider the case $B < B^*$ and $\beta < e^{-m}$. According to Theorem 6.1(c), all trajectories converge to the clipped limiting $1-$cycle; the phase portrait is presented in Figure 6.1.

To calculate the main parameters $\bar{\lambda}$, $\bar{g}$, and $\bar{x}$, we need the following quantities and functions.

- Starting point of the cycle, i.e., the minimal value of w in Figure 6.1:

$$w_0 = \mu T \beta e^{m(S_{CD}+1)}, \tag{6.12}$$

where S_{CD} is the single positive solution of

$$\mu T(e^{mS_{CD}} + e^{-S_{CD}}m) - (m+1)(\mu T + B) = 0. \tag{6.13}$$

-Duration of the cycle:

$$T_{cycle} = \frac{T}{m}\ln\frac{1}{\beta} + \frac{B}{\mu}\left\{\int_0^{S_{AB}} Y_{AB}(s)ds + \int_0^{S_{CD}} Y_{CD}(s)ds + 1\right\}, \tag{6.14}$$

where S_{AB} is the smaller positive solution of

$$0 = w_0 e^{mS_{AB}} - \mu T(m+1) + e^{-S_{AB}}[(m+1)(B+\mu T) - w_0]; \tag{6.15}$$

here

$$Y_{AB}(s) = \frac{1}{B(m+1)}\{w_0 e^{ms} - \mu T(m+1) + e^{-s}[(B+\mu T)(m+1) - w_0]\},\ s \in [0, S_{AB}], \tag{6.16}$$

$$Y_{CD}(s) = \frac{\mu T}{B(m+1)}e^{ms} - \frac{\mu T}{B} + e^{-s}\left(\frac{\mu T m}{B(m+1)}\right),\ s \in [0, S_{CD}]. \tag{6.17}$$

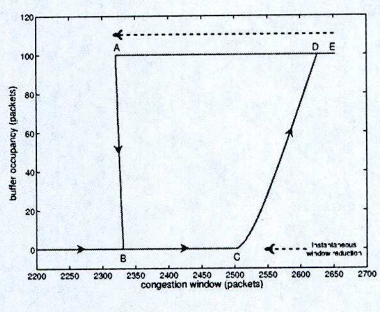

Figure 6.1: Clipped 1-cycle for Scalable TCP with μ = 1Gbps, T = 10ms, and B = 100pkts. Packet size is 4000 bits.

Proposition 6.2 *The average sending rate is given by*

$$\bar{\lambda} = \frac{w_0}{T_{cycle}m}\left(\frac{1}{\beta} - 1\right), \tag{6.18}$$

the average throughput is given by

$$\bar{g} = \frac{1}{T_{cycle}} \left\{ \frac{w_0}{m} \left(\frac{1}{\beta e^m} - 1 \right) + \mu T + B \right\}, \tag{6.19}$$

and the average amount of data in the buffer is given by

$$\begin{aligned} \bar{x} &= \frac{1}{T_{cycle}} \left\{ TB \left(\int_0^{S_{AB}} Y_{AB}(s)ds + \int_0^{S_{CD}} Y_{CD}(s)ds \right) \right. \\ &\left. + \frac{B^2}{\mu} \left(\int_0^{S_{AB}} Y_{AB}^2(s)ds + \int_0^{S_{CD}} Y_{CD}^2(s)ds \right) + B \left(T + \frac{B}{\mu} \right) \right\}. \end{aligned} \tag{6.20}$$

According to Proposition 6.1, if $B \geq B^*$ and the limiting 1-cycle is realized, then $\bar{\lambda}$ given by (6.10) is strictly greater than μ and B-independent. Thus, $(\bar{\lambda} - \mu) \not\to 0$ as $B \to \infty$. It means that in the MIMD case the rate of data loss in buffer overflow does not decrease as the buffer size increases. In contrast, in the AIMD case, we have $(\bar{\lambda} - \mu) \to 0$ as $B \to \infty$ (Avrachenkov et al (2010)). This surprising result has the following explanation. In the MIMD case, when the cycle is unclipped both the amount of data transferred over the cycle and the cycle duration are proportional to $B + \mu T$. For the parameters of the Slow Start phase of TCP New Reno with the Delayed Ack mechanism ($m = 0.5$), the expression (6.10) gives $(\bar{\lambda} - \mu)/\mu \approx 0.3$. Fortunately, the Slow Start phase switches to the Congestion Avoidance phase after the first loss is detected by triple duplicate acknowledgement (Allman et al (1999)). According to (6.10), Scalable TCP induces as little as 0.1% losses.

According to Proposition 6.2, as $B \to 0$, we have $\bar{x} \to 0$ and $\bar{g} \to \mu(\beta e^m - 1 - m)/\ln(\beta)$. In particular, in the case of Scalable TCP, we have $\bar{g} \to 0.95\mu$ as $B \to 0$. We recall from Avrachenkov et al (2010) that for AIMD, when the packet size is small in comparison with the BDP (Bandwidth Delay Product) μT, we have $\bar{g} \to \mu(1 + \beta)/2$ as $B \to 0$. Thus, the Congestion Avoidance phase of TCP New Reno with $\beta = 0.5$ has the worse link utilization of 0.75μ than that of Scalable TCP with $\beta = 0.875$ (0.95μ) when the buffer size is small. It turns out that this difference mostly comes from different values of β. In fact, one can easily check that $\mu(\beta e^m - 1 - m)/\ln(\beta) = \mu(1 + \beta)/2 + o(1 - \beta)$ and consequently, if one chooses the same value of β close to one for AIMD and MIMD, the link utilization would be the same for the two congestion control mechanisms for small buffer sizes.

6.5 Simulation results

An important claim: As mentioned earlier, this research was done in collaboration with colleagues from France. In particular, it should be made clear that the simulation results presented in this section belong to Dr. Urtzi Ayesta (LAAS-CNRS) during our collaboration.

We perform network simulations with the help of NS-2, the widely used open-source network simulator. We consider the following benchmark example of a TCP/IP network with a single bottleneck link. The topology may for instance represent an access network. The capacity of the bottleneck link is denoted by μ and its propagation delay is denoted by d. We will consider several choices for the values of μ and d. The packet size is 500bytes = 4000bits. When we simulate a scenario with multiple connections, we will assume that each connection is connected to the bottleneck link via its own access link. The capacities of the access links are supposed to be large enough so that they do not hinder the traffic.

We consider the MIMD control strategy with $m = 0.01$ and $\beta = 0.875$, that is, the standard values for Scalable TCP.

6.5.1 Impact of the buffer size on the link utilization

We first study how the utilization depends on the buffer size. We consider the values $\mu = 1$Gbps = 1 Gigabit per second and $d = 5$ms (thus $T = 2d = 10$ms).

In Figure 6.2, based on our analytical results, we plot the value of B^* (equation (6.5)) as a function of m. We observe from Figure 6.2 that for $m = 0.01$, the value of B^* is approximately 230 packets (the packet size is 4000 bits).

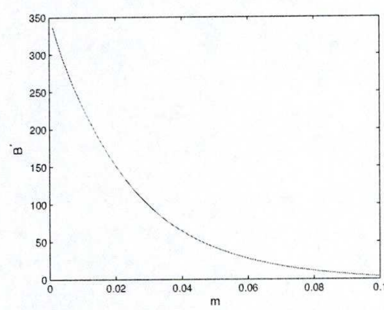

Figure 6.2: B^* (in packets) as a function of m for Scalable TCP with $\mu = 1$Gbps, $T = 10$ms, and $\beta = 0.875$.

We investigate the impact of the buffer size on the link utilization. From Theorem 6.1 it follows that according to the fluid model, $B^* = 230$ packets is the minimum buffer size such as the link is utilized at 100%. Note that the BDP for these values is equal to 2500 packets. According to the well known rule of thumb for AIMD connections (Villamizar and Song (1994)), the minimum buffer size that guarantees 100% utilization is 2500 packets.

Our fluid model predicts that for MIMD, the minimum buffer is much smaller (230 in this example). In Figure 6.3 we provide the utilization of the link for several values of the buffer size. We note that in the simulation the minimum buffer size where we observe 100% utilization is 450 packets. We note that the utilization when the buffer size is 230 packets is already quite high since it is very close to 99%. Clearly our fluid model predicts a much smaller value, which can be explained by the fact that the simulated traffic is not as smooth as it is in the fluid model. However we note that the fluid model estimation for B^* is of the same order as the optimal value obtained via simulations when comparing it with the BDP rule-of-thumb for AIMD given in Villamizar and Song (1994).

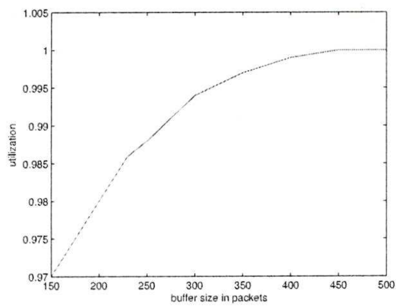

Figure 6.3: Utilization against buffer size

6.5.2 Trajectories of the dynamical systems

We simulate now the evolution in time of the congestion window, the buffer occupancy and the sending rate. We consider the same example as above, namely, $\mu = 1\text{Gbps} = 1$ Gigabit per second and $d = 5\text{ms}$ (thus $T = 2d = 10\text{ms}$). The packet size is 4000 bits. We consider again Scalable TCP, that is, $m = 0.01$ and $\beta = 0.875$.

In Figures 6.4 and 6.5 we depict the curves of $x(t)$, $w(t)$ and $\lambda(t)$ for $B = 230$ and $B = 500$, respectively. As predicted by Theorem 6.1, for $B = 230$, the cycle is critical, and the link is utilized at 100%. For $B = 500$, the cycle is unclipped and the buffer is never empty. For $B = 100$, we plot the phase portrait of a clipped cycle in the plane (w,x) in Figure 6.1 for illustrative means.

The figures for sending rate $\lambda(t)$ might appear a bit odd from the first glance. However, the flat part with the steep increasing part following it can be understood in the following way. Consider the derivative of λ with respect to t (corresponding to the part before x reaches buffer size B). Based on equations (6.1), (6.2) and (6.4), one can easily show that $\frac{d\lambda}{dt} = \frac{\lambda(m+1-\frac{\lambda}{\mu})}{T+\frac{x}{\mu}}$. Now, focusing on the numerator, clearly, $\frac{d\lambda}{dt} = 0$ when $\lambda = \mu(m+1)$, as confirmed also by the figures. Say $\lambda(\hat{t}) = \mu(m+1)$. After this point $\hat{t}$, we have a sliding mode, since $\lambda > \mu(m+1) \Rightarrow \frac{d\lambda}{dt} < 0$ and $\lambda < \mu(m+1) \Rightarrow \frac{d\lambda}{dt} > 0$. This sliding mode explains the flat part. On the other hand, this motion is up

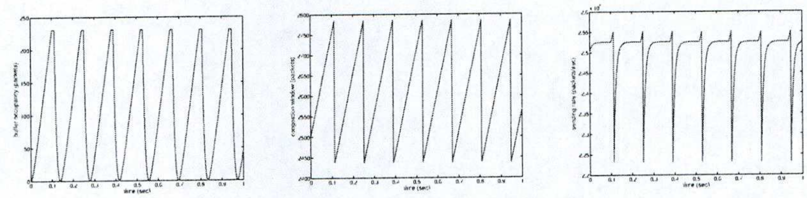

Figure 6.4: Evolution in time of the buffer occupancy, congestion window and sending rate for Scalable TCP with $\mu = 1$Gbps, $T = 10$ms, and $B = 230$pkts.

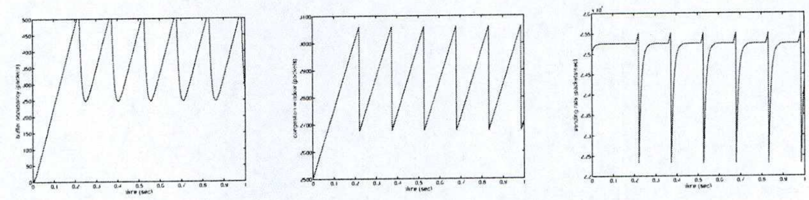

Figure 6.5: Evolution in time of the buffer occupancy, congestion window and sending rate for Scalable TCP with $\mu = 1$Gbps, $T = 10$ms, and $B = 500$pkts.

to the point when x reaches B. Then as far as x stays there, $\frac{d\lambda}{dt} = \frac{m\lambda}{T+\frac{B}{\mu}}$, explaining the steep increasing part after the flat part.

6.5.3 Pareto set

Now we compare the numerical Pareto Set with the expressions for $\bar{\lambda}$ and $\bar{g}$ given in Propositions 6.1 and 6.2. We consider AIMD (New Reno version (Allman et al (1999))) and MIMD connections. In the case of AIMD we will obtain the Pareto Set for several values of number of persistent connections, whereas for MIMD we will only consider one. We recall that several symmetric synchronized MIMD connections are equivalent to a single MIMD connection.

Let N denote the number of persistent connections in the simulation. We will assume that each connection is connected to the bottleneck link via its own access link. The capacities of the N access links leading to the bottleneck link are supposed to be large enough (or the load on each access link is small enough) so that they do not hinder the traffic. For each of these N links, the delay and capacity are $d_i = 1$ms and $\mu_i = 1000$Mbps, respectively. The fact that the delays in the access links are the same implies that the TCP connections will be synchronized.

We consider the following values for the bottleneck link: capacity is $\mu = 100$Mbps, bottleneck link propagation delay $d = 1$ms, the access link capacity and delay are 1000Mbps and 1ms, respectively. Thus $T = 2(d + d_i) = 0.004$ sec.

In Figure 6.6 we depict the Pareto set for the cases of AIMD with $N = 2$, $N = 5$

and $N = 20$ connections, and MIMD with just one connection. The qualitative shape of the curves agrees with what our model predicts. In particular, MIMD achieves the full link utilization with a much smaller buffer size than in the case of AIMD. We

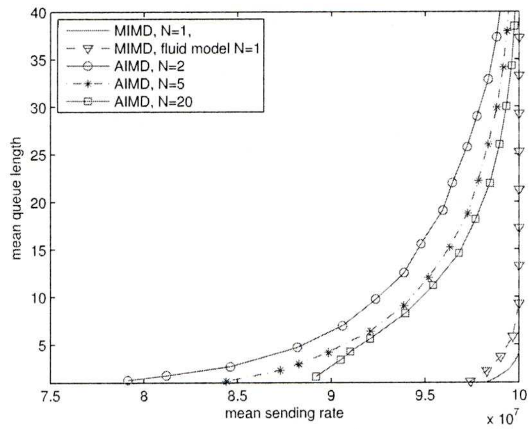

Figure 6.6: The trade-off curves for AIMD ($N = 2, N = 5, N = 20, M = 1$ packet = 500 bytes) and MIMD ($N = 1, m = 0.01$), $\beta = 0.875, T = 0.004$ sec, $\mu = 1000$ Mbps.

also display the theoretical trade-off curve for the mathematical fluid model as given in Propositions 6.1, 6.2. It turns to be close to the curve coming from simulations. However, when comparing the results obtained from the analytical model and from simulations we have observed some differences. For example, when the buffer size is zero, the simulated average sending rate is smaller than the one obtained with the fluid model. Similarly, in the simulated scenario the minimal buffer size that guarantees the full utilization of the link is larger than the one predicted by the fluid model. These differences can be explained by the fact that the traffic in the simulations is not as smooth as the fluid model that we have used.

6.6 Conclusion

We have analyzed a hybrid model for the interaction between the MIMD congestion control mechanism and a Drop Tail Internet router buffer. The present hybrid model is a significant extension of the model in Ayesta et al (2008). The present model allows us to study the impact of the time-varying Round Trip Times on the system performance. We have obtained conditions for the absence of multiple reductions of the congestion window within one congestion cycle. It turns out that these conditions are violated in the Slow Start phase of TCP New Reno without the Delayed Ack mechanism. Therefore, it is indeed recommended to use the Delayed Ack mechanism in the Slow Start phase. Fortunately, the obtained conditions are satisfied by the parameters of Scalable TCP. For Scalable TCP, we construct the Pareto set that allows us to choose a buffer size which achieves a trade off between high link utilization and small queueing delays.

6.7 Proof of main statements

The Appendix is organized as follows. Firstly, we prove a series of Lemmas and then use them to prove Theorem 6.1. The proofs of Propositions 6.1, 6.2 come at the end.

To make the model more tractable, we change the time scale and the variables as follows.

$$ds = \frac{dt}{T + x(t)/\mu}, \; y = x/B, \text{ and } v = w/B.$$

Then

$$\frac{dv}{ds} = \frac{dv}{dw}\frac{dw}{dt}\frac{dt}{ds} = \frac{mw(s)}{B} = mv(s), \tag{6.21}$$

and

$$\frac{dy}{ds} = \begin{cases} \frac{dy}{dx}\frac{dx}{dt}\frac{dt}{ds} = v(s) - q - y(s) & \text{if } 0 < y(s) < 1, \text{ or } y(s) = 0 \text{ and } v(s) > q, \\ & \text{or } y(s) = 1, \text{ and } v(s) \leq q+1 \\ 0 & \text{otherwise,} \end{cases} \tag{6.22}$$

where we have put $q = \frac{\mu T}{B}$, which is a positive constant. Now everything is in the new time scale. Let s^* be the time moment in the new time scale when the state of the system reaches 1. That is, $y(s^*) = 1$. Then the impulsive control (6.3) can now be written as

$$v(s^* + 1 + 0) = \beta^k v(s^* + 1 - 0), \tag{6.23}$$

where $k = \min\{i = 1, 2, \cdots : \beta^i v(s^* + 1 - 0) < q + 1\}$, and we notice that the time delay δ has been standardized in the new time scale. With the new variables and time scale, we see when the buffer is filled full, $y(s^*)$ reaches 1, after 1 RTT, the congestion signal is received leading to a multiplicative reduction on $v(s^*)$ with a factor β^k, where k is just defined above. Note, reducing v below $q + 1 = \frac{\mu T}{B} + 1$ exactly corresponds to reducing the instantaneous sending rate defined as $\frac{v(^*)B}{T + \frac{\mu}{B}}$ below the capacity μ.

If we ignore the non-negativity constraint on variable y, then one can solve (6.21) and (6.22) for $v(s)$ and $y(s)$ with initial conditions $v(0) = v_0$ and $y(0) = y_0$ respectively, and obtain

$$v(s) = v_0 e^{ms} \tag{6.24}$$

$$y(s) = \frac{v_0}{m+1}e^{ms} - q + e^{-s}\left(y_0 + q - \frac{v_0}{m+1}\right), \tag{6.25}$$

and the existence and uniqueness of the above two solutions follow from the initial value problems of ordinary differential equations.

With the new variables in the new time scale, we give a corresponding version of Definition 6.1 as follows.

Definition 6.1' *A cycle is defined as the trajectory starting with the initial point (a particular $V_0 \in [\beta(q+1), q+1)$, $y(0) = y_0 = 1$) at $s = 0$, and reaching the same point for the first time at some $S+1 \geq 1$. And $S+1$ is called the duration of the cycle. A cycle with $y(s)$ staying at zero for a positive time interval is called clipped. Otherwise it is unclipped. In particular, a cycle with $y(s)$ staying at 0 at a single time moment is called critical, and it could be referred to as an unclipped cycle. In addition, cycles, possibly with more than one instant jump though, are called simple, if they have only one loop. Otherwise, they are called complicated.*

In what follows, expression "unconstrained case" means that we ignore the non-negativity constraint on variable y. Expression "general case" means that we impose constraint $y \geq 0$. Under "trajectory" we mean the phase portrait $y(s)$ against $v(s)$: see Figure 6.1.

Lemma 6.1 *In the unconstrained case, 1-cycle exists iff $\beta e^m < 1$.*

Proof. Consider the unconstrained case. A 1−cycle exists iff there exists a nonnegative V_0 such that $V_0 \in [\beta(q+1), q+1)$, and it solves

$$\beta V_0 e^{m(S+1)} = V_0 \tag{6.26}$$

$$1 = \frac{V_0}{m+1} e^{mS} - q + e^{-S}\left(1 + q - \frac{V_0}{m+1}\right), \tag{6.27}$$

where we have already put $y_0 = 1$.

Firstly one can check the existence of a solution to equations (6.26) (6.27). Indeed from (6.26) we have

$$S = \frac{1}{m} \ln \frac{1}{\beta} - 1, \tag{6.28}$$

so that (6.27) results in

$$V_0 = \frac{(1+q)e^m \beta \left(1 - e\beta^{\frac{1}{m}}\right)(m+1)}{1 - e^{m+1}\beta^{\left(1+\frac{1}{m}\right)}}. \tag{6.29}$$

Secondly, one can check that V_0 given by equation (6.29) is in the interval $[\beta(q+1), q+1)$, provided $1+q > 0$ and $\beta \in (0, e^{-m})$. The latter condition is necessary and sufficient for the presented reasoning to hold. To be exact, notice that $V_0 < q+1$ iff for $\beta \in (0, e^{-m})$, the auxiliary function

$$f(\beta) = (m+1)e^m \beta \left(1 - e\beta^{\frac{1}{m}}\right) + e^{m+1}\beta^{\left(1+\frac{1}{m}\right)} = e^m \beta \left(m - me\beta^{\frac{1}{m}} + 1\right) < 1.$$

And that is true. Indeed, one can check $f(0) = 0 < 1$, $f(e^{-m}) = 1$, and

$$f'(\beta) = e^m \left[m\left(1 - e\beta^{1/m}\right) + 1 - e\beta^{1/m}\right] > 0.$$

Similarly, it is noted that $V_0 \geq \beta(q+1)$ iff for $\beta \in (0, e^{-m})$, the auxiliary function

$$g(\beta) = (m+1)e^m\left(1 - e\beta^{\frac{1}{m}}\right) + e^{m+1}\beta^{1+\frac{1}{m}} \geq 1.$$

And that is true, because $g(0) = (m+1)e^m > 1$, $g(e^{-m}) = 1$, and

$$g'(\beta) = -e^{m+1}\left[\beta^{1/m}\left(\frac{1}{\beta} - 1\right) + \frac{1}{m}\beta^{1/m}\left(\frac{1}{\beta} - 1\right)\right] < 0.$$

■

In what follows, it is assumed that $e^m\beta < 1$.

Lemma 6.2 *Consider the unconstrained case. Starting from an arbitrary initial state $v_0 \in (0, 1+q)$, $y_0 = 1$, component $y(s)$ attains its single minimum at the moment*

$$s_1(v_0) = \frac{1}{m+1}\ln\frac{(1+q)(1+m) - v_0}{mv_0} > 0. \tag{6.30}$$

The value $y(s_1)$ increases with v_0 and $y(s_1) \to 1$ as $v_0 \to 1+q$.

A 1−cycle is critical for a single nonnegative value of q given by

$$q^* = \frac{(m+1)e^m\left(1 - e\beta^{\frac{1}{m}}\right)\beta e^{ms_1^*}}{1 - e^{m+1}\beta^{\frac{m+1}{m}} - (m+1)e^m\left(1 - e\beta^{\frac{1}{m}}\right)\beta e^{ms_1^*}}, \tag{6.31}$$

where $s_1^* = s_1(V_0^*) = \frac{1}{m+1}\ln\left(\frac{1-e^m\beta}{\beta e^m m\left(1-e\beta^{\frac{1}{m}}\right)}\right)$, and V_0^* is given by (6.29) with $q = q^*$.

In the general case (if we impose constraint $y \geq 0$) the 1−cycle is unclipped iff $q \leq q^*$.

Proof. According to equations (6.24) (6.25), we have the following equations satisfied by $s_1(v_0)$:

$$\begin{cases} v(s_1) = v_0 e^{ms_1} \\ y(s_1) = \frac{v_0}{m+1}e^{ms_1} - q + e^{-s_1}(1 + q - \frac{v_0}{m+1}) \\ y(s_1) = v(s_1) - q. \end{cases}$$

Indeed, $s_1(v_0)$ is given by the solution to the equation:

$$
\begin{aligned}
& \frac{v_0}{m+1}e^{ms_1(v_0)} - q + e^{-s_1(v_0)}(1+q-\frac{v_0}{m+1}) = v_0 e^{ms_1(v_0)} - q \\
\Leftrightarrow\quad & v_0 e^{ms_1(v_0)} + e^{-s_1(v_0)}((1+q)(1+m)-v_0) = v_0(m+1)e^{ms_1(v_0)} \\
\Leftrightarrow\quad & m v_0 e^{ms_1(v_0)} = e^{-s_1(v_0)}((1+q)(1+m)-v_0) \\
\Leftrightarrow\quad & e^{(m+1)s_1(v_0)} = \frac{(1+q)(1+m)-v_0}{mv_0},
\end{aligned}
$$

which gives

$$
s_1(v_0) = \frac{1}{m+1}\ln\frac{(1+q)(1+m)-v_0}{mv_0} > 0,
$$

because $v_0 < 1+q$.

Put $s_1(v_0)$ given by (6.30) in (6.25) and obtain

$$
\begin{aligned}
y(s_1(v_0)) &= v_0\left(\frac{(1+q)(1+m)-v_0}{mv_0}\right)^{\frac{m}{m+1}} - q \\
&= \left\{\frac{((1+q)(1+m)-v_0)v_0 v_0^{\frac{1}{m}}}{mv_0}\right\}^{\frac{m}{m+1}} - q \\
&= \left\{\frac{((1+q)(1+m)-v_0)v_0^{\frac{1}{m}}}{m}\right\}^{\frac{m}{m+1}} - q \\
&= \left\{\frac{(1+q)(1+m)v_0^{\frac{1}{m}} - v_0^{\frac{1+m}{m}}}{m}\right\}^{\frac{m}{m+1}} - q.
\end{aligned}
$$

One can easily check that $y(s_1(v_0)) \to 1$ as $v_0 \to 1+q$. For $\frac{dy(s_1(v_0))}{dv_0}$, we have

$$
\begin{aligned}
\frac{dy(s_1(v_0))}{dv_0} &= \frac{m}{m+1}\left\{\frac{(1+q)(1+m)v_0^{\frac{1}{m}} - v_0^{\frac{1+m}{m}}}{m}\right\}^{\frac{-1}{m+1}} \\
&\quad \left\{\frac{(1+q)(1+m)}{m^2}v_0^{\frac{1}{m}-1} - \frac{1+m}{m^2}v_0^{\frac{1}{m}}\right\} \\
&= \frac{m}{m+1}\left\{\frac{(1+q)(1+m)v_0^{\frac{1}{m}} - v_0^{\frac{1+m}{m}}}{m}\right\}^{\frac{-1}{m+1}} \\
&\quad \left\{\frac{1+m}{m^2}v_0^{\frac{1}{m}-1}(1+q-v_0)\right\} > 0. \text{ (Recall } v_0 < 1+q.\text{)}
\end{aligned}
$$

Let us fix an arbitrary $q > 0$ and consider the corresponding simple $1-$cycle with the corresponding value of V_0 defined in (6.29). Now

$$s_1(V_0) = \frac{1}{m+1}\ln\frac{1-e^m\beta}{me^m\beta(1-e\beta^{1/m})} = s_1^* \tag{6.32}$$

and according to (6.22) and (6.24)

$$y(s_1(V_0)) = v(s_1) - q = V_0 e^{ms_1} - q.$$

Since s_1^* is $q-$independent, $y(s_1(V_0))$ is a linear function of q. Let us show that it decreases with q.

Indeed, if $q \to 0$ then $y(s_1(V_0))$ has a positive limit. When q increases, $y(s_1(V_0))$ becomes negative. To see this, notice that at the beginning of the cycle, starting from $v(0) = V_0 < 1+q$, $y(0) = 1$, component y decreases. Moreover,

$$\begin{aligned} \frac{d}{ds}\left(\frac{y(s)}{q}\right)\Big|_{s=0} &= \frac{V_0 m e^{ms}}{(m+1)q} - \frac{e^{-s}(1+q-\frac{V_0}{1+m})}{q}\Big|_{s=0} \\ &= \frac{V_0}{q} - 1 - \frac{1}{q} \to \frac{e^m\beta(1-e\beta^{1/m})(m+1)}{1-e^{m+1}\beta^{1+1/m}} - 1 \end{aligned}$$

as $q \to \infty$. And the latter expression is negative because the $\lim_{\beta\to 0} e^m\beta(1-e\beta^{1/m})(m+1) - 1 + e^{m+1}\beta^{1+1/m} = -1$, $\lim_{\beta\to e^{-m}} e^m\beta(1-e\beta^{1/m})(m+1) - 1 + e^{m+1}\beta^{1+1/m} = 0$, and $\frac{d}{d\beta}(e^m\beta(1-e\beta^{1/m})(m+1) - 1 + e^{m+1}\beta^{1+1/m}) = e^m(m+1)(1-e\beta^{1/m}) > 0$ for $\beta \in (0, e^{-m})$. Therefore, $\frac{y(s)}{q}$ decreases with time s, when s is small, at large values of q, starting from initial value $\frac{y_0}{q} = \frac{1}{q}$, meaning that $\frac{y(s)}{q}$ takes negative values if q is sufficiently big, i.e. the minimal value, $y(s_1) < 0$.

Therefore, there exists a single value $q^* > 0$ such that $y(s_1(V_0)) = 0$. Clearly, the last equality holds iff

$$V_0^* e^{ms_1^*} - q^* = \frac{(1+q^*)e^m\beta(1-e\beta^{1/m})(m+1)}{1-e^{m+1}\beta^{1+1/m}} e^{ms_1^*} - q^* = 0.$$

It only remains to solve the equation obtained for q^*.

The last statement is obvious. ∎

Remark 6.2 *According to (6.29),*

$$\begin{aligned} \frac{dV_0}{d\beta} &= \frac{(1+q)e^m(m+1)[1-e(1+\frac{1}{m})\beta^{1/m} - e^{m+1}\beta^{1+1/m} + e^{m+1}(1+\frac{1}{m})\beta^{1+1/m}]}{(1-e^{m+1}\beta^{1+1/m})^2} \\ &> 0, \end{aligned}$$

if $\beta \in (0, e^{-m})$. Indeed,

$$\lim_{\beta \to 0} \left(1 - e(1+\frac{1}{m})\beta^{1/m} - e^{m+1}\beta^{1+1/m} + e^{m+1}(1+\frac{1}{m})\beta^{1+1/m} \right) = 1,$$

$$\lim_{\beta \to e^{-m}} \left(1 - e(1+\frac{1}{m})\beta^{1/m} - e^{m+1}\beta^{1+1/m} + e^{m+1}(1+\frac{1}{m})\beta^{1+1/m} \right) = 0,$$

and

$$\begin{aligned} & \frac{d}{d\beta}(1 - e(1+\frac{1}{m})\beta^{1/m} - e^{m+1}\beta^{1+1/m} + e^{m+1}(1+\frac{1}{m})\beta^{1+1/m}) \\ = & -\frac{1}{m}(1+1/m)e\beta^{1/m-1} - (1+1/m)e^{m+1}\beta^{1/m} + (1+1/m)^2 e^{m+1}\beta^{1/m} \\ = & \; e\beta^{1/m-1}(1+1/m)\frac{1}{m}(-1+e^m\beta) < 0. \end{aligned}$$

Trajectories $(v(s), y(s))$ cannot cross when starting from different initial points $(v(0) = V_0^1, y(0) = 1)$ and $(v(0) = V_0^2, y(0) = 1)$; thus the minimal value $y(s_1(V_0))$ increases with β.

Corollary 6.1 *In the general case, where constraint $y(s) \geq 0$ is imposed, if a trajectory starting with some $v_0 \in (0, 1+q)$ is clipped, there will be some $\hat{v}_0 \in (v_0, 1+q)$, starting with which the trajectory just touches the horizontal v axis, i.e., $y(s_1(\hat{v}_0)) = 0$. Furthermore, trajectories starting with $v_0 \in [\hat{v}_0, 1+q)$ are unclipped, while those with $v_0 \in (0, \hat{v}_0)$ are clipped. As a result, if $q \leq q^*$, $V_0 \geq \hat{v}_0$, where V_0 is given by (6.29).*

Proof. Everything follows directly from the first part of Lemma 6.2, bearing in mind that increasing $y(s_1(v_0))$ is a continuous function of v_0. ■

Lemma 6.3 *As $m \to 0$, $B^* = \frac{\mu T}{\beta}(1 - \beta + m\ln m) + o(m\ln m) \to \frac{\mu T(1-\beta)}{\beta}$.*

Proof.

$$\begin{aligned} \frac{ds_1}{dm} = & \frac{1}{(m+1)^2}\left\{\left\{\frac{-\beta e^m}{1-e^m\beta} - 1 - \frac{1}{m} - \frac{e\beta^{1/m}\ln\beta}{(1-e\beta^{1/m})m^2}\right\}(m+1)\right. \\ & \left. - \left\{\ln(1-e^m\beta) - \ln\beta - m - \ln m - \ln(1-e\beta^{1/m})\right\}\right\}. \end{aligned}$$

We effectively prove

$$\begin{aligned} & \lim_{m \to 0} \frac{1}{m\ln m}\left\{\frac{1 - e^{m+1}\beta^{1+1/m}}{(m+1)e^m(1-e\beta^{1/m})e^{ms_1}} - 1 - m\ln m\right\} \\ = & \lim_{m\to 0} \frac{M_1(m) - 1 - M_2(m)}{M_2(m)} = 0 \Leftrightarrow \\ = & \lim_{m\to 0} \frac{M_1(m) - 1}{M_2(m)} = 1 \; (M_1(m) = 1 - e^{m+1}\beta^{1+\frac{1}{m}},\; M_2(m) = m\ln m), \end{aligned}$$

where L'Hôpital's rule is obviously applicable, because $\lim_{m\to 0} e^{ms_1} = 1$ as can be easily verified. Notice that

$$\frac{d(m\ln m)}{dm} = 1 + \ln m,$$

and

$$\begin{aligned}
\frac{d(M_1(m)-1)}{dm} &= \frac{1}{(m+1)^2 e^{2m}(1-e\beta^{1/m})^2 e^{2ms_1}} \Bigg\{ - \Bigg(e^{m+1}\beta^{1+1/m} \\
&\quad - \frac{e^{m+1}\beta^{1+1/m}\ln\beta}{m^2}(m+1)e^m(1-e\beta^{1/m})e^{ms_1} \Bigg) \\
&\quad -(1-e^{m+1}\beta^{1+1/m})\frac{d(m+1)e^m(1-e\beta^{1/m})e^{ms_1}}{dm} \Bigg\} \\
&= \frac{-e^{m+1}\beta^{1+1/m}(m+1)(1-\frac{\ln\beta}{m^2})(1-e\beta^{1/m})}{(m+1)^2 e^m (1-e\beta^{1/m})^2 e^{ms_1}}
\end{aligned}$$

$$\begin{aligned}
&-(1-e^{m+1}\beta^{1+1/m})((m+1)^2 e^m(1-e\beta^{1/m})^2 e^{ms_1})^{-1} \\
&\times \Bigg\{ \Bigg((2+m)(1-e\beta^{1/m}) + \frac{(m+1)e\beta^{1/m}\ln\beta}{m^2} \Bigg) \\
&+(m+1)(1-e\beta^{1/m})(s_1 + m\frac{ds_1}{dm})) \Bigg\},
\end{aligned}$$

where

$$\lim_{m\to 0} \frac{-e^{m+1}\beta^{1+1/m}(m+1)(1-\frac{\ln\beta}{m^2})(1-e\beta^{1/m})}{(m+1)^2 e^m (1-e\beta^{1/m})^2 e^{ms_1})} = 0$$

because one can easily verify that $\lim_{m\to 0} \frac{\beta^{1/m}}{m^2} = 0$.

Furthermore,

$$\lim_{m\to 0} \frac{1-e\beta^{1+1/m}}{(m+1)^2 e^m (1-e\beta^{1/m})^2 e^{ms_1}} = 1,$$

$$\lim_{m\to 0} \left((2+m)(1-e\beta^{1/m}) + \frac{(m+1)e^{m+1}\beta^{1/m}\ln\beta}{m^2} \right) = 2,$$

$$\lim_{m\to 0} \frac{(m+1)(1-e\beta^{1/m})s_1}{1+\ln m} = \lim_{m\to 0} \frac{-\ln m}{1+\ln m} = -1,$$

and

$$\begin{aligned}
&\lim_{m\to 0}(m+1)(1-e\beta^{1/m})m\frac{ds_1}{dm} \\
&= \lim_{m\to 0} m \left\{ \left\{ \frac{-\beta e^m}{1-e^m\beta} - 1 - \frac{1}{m} - \frac{e\beta^{1/m}\ln\beta}{(1-e\beta^{1/m})m^2} \right\} (m+1) \right. \\
&\quad \left. - \left\{ \ln(1-e^m\beta) - \ln\beta - m - \ln m - \ln(1-e\beta^{1/m}) \right\} \right\} = -1.
\end{aligned}$$

Finally all these collectively lead to $\lim_{m\to 0}\left(\frac{dM_1(m)-1}{dm}/\frac{dM_2(m)}{dm}\right) = 1$, as required. ■

After the continuous trajectory starting with $v(0) = v_0 < 1+q$ and $y(0) = 1$ completes, that is, the buffer is filled up and the congestion is noticed after the delay, there will be a reduction on the variable v leading to $v_1 \in [\beta(1+q), 1+q)$. Therefore, as the process proceeds, we have a sequence $\{v_i\}$. If this sequence has a limit, namely v_∞, a limiting cycle exists and will be realized.

According to (6.24) and (6.25), we introduce the following denotations (for $v < 1+q$):

$$\varphi(v) = \beta v e^{m(s+1)}, \tag{6.33}$$

where $s > 0$ solves equation

$$F(v,s) = \frac{v}{m+1}\left(e^{ms} - e^{-s}\right) + (1+q)\left(e^{-s} - 1\right) = 0. \tag{6.34}$$

In the unconstrained case, (or if an actual continuous trajectory is unclipped), if only one jump is sufficient, $v_{i+1} = \varphi(v_i)$. We shall also use the denotation $\psi(v) = \varphi(\varphi(v)) = \varphi^2(v)$ for brevity.

Note that, if the actual continuous trajectory starting from $v(0) = v$, $y(0) = 1$ is clipped then, at the next time moment s^* when $y(s^*) = 1$, $v(s^*) < ve^{ms}$ implying $v(s^* + 1 + 0) < \varphi(v)$, provided that only one jump is sufficient in the unconstrained case.

Lemma 6.4 *In the unconstrained case, starting with an arbitrary $v_0 \in [\beta e^m(1+q), 1+q)$, the limiting simple cycle exists and is of order one.*

Proof. To start off, let us check that there exists only one $s > 0$ solving (6.34) for $v \in (0, 1+q)$. To see this, notice that $\lim_{s\to 0} F(v,s) = 0$, $\lim_{s\to\infty} F(v,s) = \infty$, and $\frac{\partial F(v,s)}{\partial s} = e^{-s}\left\{\frac{v}{m+1}(me^{(m+1)s} + 1) - (1+q)\right\} = 0$ has only one single positive solution. In addition, $\lim_{s\to 0}\frac{\partial F(v,s)}{\partial s} = v - (1+q) < 0$. The above reasoning leads to the function $F(v,s)$ decreasing from zero and increasing up to infinity after a single stationary point, resulting in a single solution positive s solving $F(v,s) = 0$.

It is convenient to investigate the mapping φ defined on the closed segment $[\beta e^m(1+q), 1+q]$: $\varphi(1+q) = \beta e^m(1+q)$. (Equation (6.34) has only one zero solution for $v = 1+q$ and $\lim_{v\to 1+q-0}\varphi(v) = \beta e^m(1+q)$.)

Our proof will be done in three parts:

1. $\varphi(v)$ decreases with v. Hence $\psi(v)$ increases with v. This statement holds for all $v \in (0, 1+q)$.
2. $\varphi : [\beta e^m(1+q), 1+q] \to [\beta e^m(1+q), 1+q]$ and $\psi : [\beta e^m(1+q), 1+q] \to [\beta e^m(1+q), 1+q]$.
3. $\{\varphi^n(v)\}$ and $\{\psi^n(v)\}$ both converge to $v_\infty \in [\beta e^m(1+q), 1+q)$.

For **item 1**, according to implicit differentiation and partial differentiation,

$$\begin{aligned}\frac{d\varphi(v)}{dv} &= \beta e^{m(s+1)} + \frac{\beta v e^{m(s+1)} m\left(e^{-s}-e^{ms}\right)}{v\left(me^{ms}+e^{-s}\right)-(m+1)(q+1)e^{-s}} \\ &= \beta e^{-s}\frac{vme^{m(s+1)}+e^{m(s+1)}v-e^{m(s+1)}(m+1)(q+1)}{v\left(me^{ms}+e^{-s}\right)-(m+1)(q+1)e^{-s}} \\ &= \frac{\beta e^{-s}\left\{ve^{m(s+1)}(m+1)-e^{m(s+1)}(m+1)(q+1)\right\}}{v\left(me^{ms}+e^{-s}\right)-(m+1)(q+1)e^{-s}} \\ &= \frac{\beta e^{-s}(m+1)e^{m(s+1)}(v-(q+1))}{v\left(me^{ms}+e^{-s}\right)-(m+1)(q+1)e^{-s}},\end{aligned} \tag{6.35}$$

where the numerator of the last expression is smaller than zero for $v < 1+q$.

The denominator of the last expression equals

$$v\left(e^{ms}m+e^{-s}\right)-(m+1)(q+1)e^{-s} = \frac{(1+q)(m+1)e^{-s}}{e^{ms}-e^{-s}}G_1(s),$$

where $G_1(s) = me^{(m+1)s} - me^{ms} + 1 - e^{ms}$. (We have put in $v = \frac{(1+q)(1-e^{-s})(m+1)}{e^{ms}-e^{-s}}$, according to equation (6.34).) Finally, $G_1(s) > 0$ for $s > 0$, because $\lim_{s\to 0} G_1(s) = 0$, and $\frac{dG(s)}{ds} = me^{ms}(me^{s} - m + e^{s} - 1) > 0$.

For **item 2**, we consider $v \in [\beta e^{m}(1+q), 1+q)$. According to item 1, $\varphi(1+q)$ and $\varphi(\beta e^{m}(1+q))$ give a lower and an upper bounds for $\varphi(v)$, respectively. We then need to show that $\varphi(1+q) \geq \beta e^{m}(1+q) > \beta(1+q)$, and $\varphi(\beta e^{m}(1+q)) < 1+q$. Since $\varphi(1+q) = \beta e^{m}(1+q)$, it remains to prove that $\varphi(\beta e^{m}(1+q)) \leq 1+q$.

According to (6.34), where we put in $v = \beta e^{m}(1+q)$, we have

$$\frac{(1+q)\beta e^{m}}{m+1}\left(e^{ms}-e^{-s}\right)+(1+q)\left(e^{-s}-1\right)=0 \Leftrightarrow G_2(s,\beta,m)=0,$$

where $G_2(s,\beta,m) = \beta e^{m}\left(e^{ms}-e^{-s}\right)+(m+1)\left(e^{-s}-1\right)$. Then we observe that

$$\lim_{s\to 0} G_2(s,\beta,m) = 0,$$

$$G_2(s,\beta,m) \to \infty$$

as $s \to \infty$,

$$\frac{\partial G_2(s,\beta,m)}{\partial s} = \beta e^{m}\left(me^{ms}+e^{-s}\right)-(m+1)e^{-s},$$

and

$$\left.\frac{\partial G_2(s,\beta,m)}{\partial s}\right|_{s=0} = (\beta e^{m}-1)(m+1) < 0,$$

$$\frac{\partial G_2(s,\beta,m)}{\partial s} = 0 \Leftrightarrow s = \frac{1}{m+1}\ln\frac{m+1-\beta e^{m}}{\beta e^{m} m} > 0,$$

and
$$\frac{\partial G_2(s,\beta,m)}{\partial s} \longrightarrow \infty$$
as $s \longrightarrow \infty$. Hence function $G_2(s,\beta,m)$ firstly decreases with respect to s from zero and then increases up to infinity after the single minimum point, giving a single positive solution s solving (6.34) with $v = \beta e^m(1+q)$.

Clearly, $\varphi(\beta e^m(1+q)) = \beta^2(1+q)e^m e^{m(s+1)}$, where s solves (6.34) with $v = \beta e^m(1+q)$. Define the increasing (with respect to s) auxiliary function $G_3(s) = \beta^2(1+q)e^m e^{m(s+1)}$. We aim to show that for $\hat{s}$ such that $G_3(\hat{s}) = 1+q$, i.e., $\hat{s}(\beta,m) = \frac{2}{m}\ln\frac{1}{\beta} - 2$, $G_2(\hat{s}(\beta,m),\beta,m) > 0$. That would say, $\hat{s}(\beta,m)$ is greater than the solution of (6.34) with $v = \beta e^m(1+q)$, and $\varphi(\beta e^m(1+q)) < 1+q$.

We have
$$G_2(\hat{s}(\beta,m),\beta,m) = (e^m\beta)^{-1} - \beta e^m\left(e\beta^{\frac{1}{m}}\right)^2 + (m+1)\left(\left(e\beta^{\frac{1}{m}}\right)^2 - 1\right) = \hat{G}_2(\beta,m).$$

Observe that $\hat{G}_2(\beta,m) \longrightarrow \infty$ as $\beta \longrightarrow 0$ and $\hat{G}_2(\beta,m) \longrightarrow 0$ as $\beta \longrightarrow e^{-m}$.

Furthermore,
$$\begin{aligned}\frac{\partial \hat{G}_2(\beta,m)}{\partial \beta} &= e^{-m}\beta^{-2}\left(-1 - \frac{m+2}{m}e^{2m+2}\beta^{\frac{2m+2}{m}} + \frac{2(m+1)}{m}e^{m+2}\beta^{\frac{2+m}{m}}\right)\\ &= e^{-m}\beta^{-2}\hat{G}_3(\beta,m),\end{aligned}$$

where $\hat{G}_3(\beta,m) = -1 - \frac{m+2}{m}e^{2m+2}\beta^{\frac{2m+2}{m}} + \frac{2(m+1)}{m}e^{m+2}\beta^{\frac{2+m}{m}} < 0$ for $\beta \in (0,e^{-m})$, because $\lim_{\beta\to 0}\hat{G}_3(\beta,m) = -1 < 0$, $\hat{G}_3(\beta,m) \longrightarrow -1 - \frac{m+2}{m} + \frac{2(m+1)}{m} = 0$ as $\beta \longrightarrow e^{-m}$, and
$$\begin{aligned}\frac{\partial \hat{G}_3(\beta,m)}{\partial \beta} &= -\frac{m+2}{m}e^{2m+2}\beta^{\frac{2+m}{m}}\frac{2+2m}{m} + \frac{2m+2}{m}e^{m+2}\frac{2+m}{m}\beta^{\frac{2}{m}}\\ &= \frac{2m+2}{m}e^{m+2}\frac{2+m}{m}\beta^{\frac{2}{m}}(1-\beta e^m) > 0.\end{aligned}$$

Therefore, $\frac{\partial \hat{G}_2(\beta,m)}{\partial \beta} < 0 \Rightarrow \hat{G}_2(\beta,m) > 0 \Leftrightarrow G_2(\hat{s}(\beta,m),\beta,m) > 0$, as required.

It follows from item 1 that starting with an arbitrary $v \in [\beta e^m(1+q), 1+q)$, $\{\psi^n(v)\}$ is a monotonic sequence. It follows from item 2 that the sequence $\{\psi^n(v)\}$ is bounded in the closed interval $[\beta e^m(1+q), 1+q]$. Hence, $\psi^n(v_0) \longrightarrow v_\infty = \psi(v_\infty) \in [\beta e^m(1+q), 1+q]$ as $n \longrightarrow \infty$. It also follows from item 2 that with $e^m\beta < 1$, exactly one jump is enough, starting with $v \in [\beta e^m(1+q), 1+q)$.

For **item 3**, assume $\varphi(v_\infty) = v'_\infty \neq v_\infty$. Let S_2 and S_3 be such that
$$0 = \frac{v'_\infty}{m+1}\left(e^{mS_2} - e^{-S_2}\right) + (1+q)\left(e^{-S_2} - 1\right) \tag{6.36}$$
$$0 = \frac{v_\infty}{m+1}\left(e^{mS_3} - e^{-S_3}\right) + (1+q)\left(e^{-S_3} - 1\right). \tag{6.37}$$

Then S_2+1 and S_3+1 are the durations of continuous trajectories starting with v'_∞ and v_∞, respectively. (See (6.34).) Then by the definition of v_∞,

$$v_\infty = \beta\beta v_\infty e^{m(S_3+1)} e^{m(S_2+1)},$$

leading to

$$\beta = e^{-\frac{1}{2}m(S_2+S_3+2)}. \tag{6.38}$$

By the definition of v'_∞, we have $v'_\infty = \beta v_\infty e^{m(S_3+1)}$. But from (6.36) we have in parallel $v'_\infty = \frac{(1+q)(m+1)(1-e^{-S_2})}{e^{mS_2}-e^{-S_2}}$. Equating these two expressions, substituting the expression for v_∞ coming from (6.37), and using formula (6.38), we have that

$$\frac{(1+q)(1+m)(1-e^{-S_2})}{e^{mS_2}-e^{-S_2}} = \frac{(1+q)(1+m)(1-e^{-S_3})e^{-\frac{1}{2}(S_2+S_3+2)m}e^{m(S_3+1)}}{e^{mS_3}-e^{-S_3}}$$

$$\Leftrightarrow \quad \frac{(1-e^{-S_2})e^{\frac{1}{2}mS_2}}{e^{mS_2}-e^{-S_2}} = \frac{(1-e^{-S_3})e^{\frac{1}{2}mS_3}}{e^{mS_3}-e^{-S_3}}. \tag{6.39}$$

Now define the auxiliary function $G_3(z) = \frac{(1-e^{-z})e^{\frac{1}{2}mz}}{e^{mz}-e^{-z}} = \frac{e^{\frac{1}{2}mz}-e^{(\frac{1}{2}m-1)z}}{e^{mz}-e^{-z}}$. Then

$$\begin{aligned}
\frac{dG_3(z)}{dz} &= \left\{\left(\frac{1}{2}me^{\frac{1}{2}mz} - \left(\frac{1}{2}m-1\right)e^{(\frac{1}{2}m-1)z}\right)\left(e^{mz}-e^{-z}\right)\right. \\
&\quad \left. - \left(e^{\frac{1}{2}mz} - e^{(\frac{1}{2}m-1)z}\right)\left(me^{mz}+e^{-z}\right)\right\}\left(e^{mz}-e^{-z}\right)^{-2} \\
&= \frac{G_4(z)}{(e^{mz}-e^{-z})^2}.
\end{aligned}$$

After some rearrangements we obtain

$$\begin{aligned}
G_4(z) &= e^{-2z}e^{\frac{1}{2}mz}\left\{\frac{1}{2}m - \frac{1}{2}me^{(m+2)z} + \frac{1}{2}me^{(m+1)z}\right. \\
&\quad \left. -\frac{1}{2}me^{z} + e^{(m+1)z} - e^{z}\right\} \\
&= e^{-2z}e^{\frac{1}{2}mz}G_5(z).
\end{aligned}$$

It is easy to see that $\lim_{z\to 0} G_5(z) = 0$, and $G_5(z) < 0$ as $z \to \infty$. For $\frac{dG_5(z)}{dz}$,

$$\begin{aligned}
\frac{dG_5(z)}{dz} &= -\frac{1}{2}m(m+2)e^{(m+2)z} + \frac{1}{2}m(m+1)e^{(m+1)z} \\
&\quad -\frac{1}{2}me^{z} + (m+1)e^{(m+1)z} - e^{z} \\
&= e^{z}\left\{-\frac{1}{2}m(m+2)e^{(m+1)z} + \frac{1}{2}m(m+1)e^{mz}\right. \\
&\quad \left. -\frac{1}{2}m + (m+1)e^{mz} - 1\right\} \\
&= e^{z}G_6(z).
\end{aligned}$$

Observe that $\lim_{z\to 0} G_6(z) = 0$ and $G_6(z) < 0$ as $z \to \infty$. Furthermore,

$$\begin{aligned}
\frac{dG_6(z)}{dz} &= -\frac{1}{2}m(m+2)(m+1)e^{(m+1)z} + \frac{1}{2}m^2(m+1)e^{mz} + m(m+1)e^{mz} \\
&= e^{mz}\left\{-\frac{1}{2}m(m+2)(m+1)e^{z} + \frac{1}{2}m^2(m+1) + m(m+1)\right\} \\
&= e^{mz}G_7(z) < 0,
\end{aligned}$$

because $G_7(0) = 0$, $G_7(z) < 0$ as $z \to \infty$, and obviously $\frac{dG_7(z)}{dz} < 0$. Therefore, function $\frac{(1-e^{-z})e^{\frac{1}{2}mz}}{e^{mz}-e^{-z}}$ strictly decreases if $z > 0$. Hence $S_2 = S_3$ and $v'_\infty = v_\infty < 1+q$ because $\varphi(1+q) = \beta e^{m}(1+q) \neq 1+q$. Consequently, $\varphi^n(v_0) \to v_\infty$ as $n \to \infty$. ■

Remark 6.3 *It follows from* ***item 3*** *in the proof of Lemma 6.4 that in the unconstrained case, starting with an arbitrary $v_0 \in [\beta e^{m}(1+q), 1+q)$, complicated cycles cannot be realized, and v_∞ coincides with V_0 given by (6.29).*

Corollary 6.2 $-1 < \left.\frac{d\varphi(v_0)}{dv_0}\right|_{V_0} < 0$, *so that the mapping φ is a contraction in a neighbourhood of the stable point V_0.*

Proof. By putting in V_0 given by (6.29) and S given by (6.28) into (6.35), we have $\left.\frac{d\varphi(v)}{dv}\right|_{V_0} = \frac{e\beta^{1/m}[(m+1)e^{m}\beta - 1 - me^{m+1}\beta^{1+1/m}]}{(1-e\beta^{1/m})m + e^{m+1}\beta^{1+1/m} - e\beta^{1/m}}$. We already know that $\frac{d\varphi(v)}{dv} < 0$ (see item 1 above). Hence we just need to prove that $\left.\frac{d\varphi(v_0)}{dv_0}\right|_{V_0} > -1 \Leftrightarrow P_1(\beta, m) > 0$, where $P_1(\beta, m) = (m+1)e^{m}\beta - 2 - me^{m+1}\beta^{\frac{1+m}{m}} + e^{m}\beta - m + m(e\beta^{\frac{1}{m}})^{-1}$. But $P_1(\beta, m) \to \infty$ as $\beta \to 0$, $P_1(\beta, m) \to 0$ as $\beta \to e^{-m}$, $\frac{\partial P_1(\beta,m)}{\partial\beta} = (m+1)e^{m} - (1+m)e^{m+1}\beta^{\frac{1}{m}} + e^{m} - e^{-1}\beta^{\frac{-1-m}{m}}$, $\frac{\partial P_1(\beta,m)}{\partial\beta} \to -\infty$ as $\beta \to 0$, $\frac{\partial P_1(\beta,m)}{\partial\beta} \to 0$ as $\beta \to e^{-m}$, and $\frac{\partial P_2(\beta,m)}{\partial\beta} = -e^{m+1}\frac{m+1}{m}\beta^{\frac{1}{m}-1} + e^{-1}\frac{m+1}{m}\beta^{\frac{-2m-1}{m}} = \frac{m+1}{m}e^{-1}\beta^{\frac{-2m-1}{m}}\left(1 - e^{m+2}\beta^{\frac{m+2}{m}}\right) > 0$ altogether result in that function $P_1(\beta, m)$ monotonically decreases from ∞ to 0 for $\beta \in (0, e^{-m})$.
■

Corollary 6.3 *In the unconstrained case, let $v_0 \in [\beta e^{m}(1+q), 1+q)$. Then $\forall i \in \{0, 1, 2, \ldots\}$, $v_i \in [\beta e^{m}(1+q), 1+q)$, and $v_{i+2} \in [\min(v_i, v_{i+1}), \max(v_i, v_{i+1})]$.*

Proof. The first statement follows from the proof of Lemma 6.4.

Without loss of generality, we can put $i = 0$. That is, we aim to show that $v_2 \in [\min(v_0, v_1), \max(v_0, v_1)]$. According to **item 2** in the proof of Lemma 6.4, $v_i = \varphi(v_{i-1})$. Consider the case $v_0 > v_1$. Automatically we have $v_2 > v_1$, since φ is decreasing. Then there are two possibilities about the relationship between v_0, v_1, and v_2:

1. $v_2 > v_0 > v_1$.

2. $v_2 \in [v_1, v_0]$.

Suppose the first possibility is true, that is, $v_2 > v_0 > v_1$. We aim to show by induction that in this case, $v_{2i+2} > v_{2i} > \cdots > v_2 > v_0 > v_1 > \cdots > v_{2i+1}$, $\forall i \in \{0, 1, 2, \dots\}$. These inequalities hold for $i = 0$. Suppose they hold for some $i \geq 0$. Consider the case $i+1$. From the induction supposition we have $v_{2i+2} > v_{2i}$. Therefore $v_{2i+1} > v_{2i+3}$, and $v_{2i+4} > v_{2i+2}$. Hence sequence $\{v_i\}$ does not converge which contradicts Lemma 6.4.

Therefore, the first possibility is false. And $v_2 \in [v_1, v_0]$ holds automatically, as we want. Exactly in the same manner, one can show that in the case $v_0 < v_1$, $v_2 \in [v_0, v_1]$. And the case $v_0 = v_1$ is trivial. Hence, $v_2 \in [\min(v_0, v_1), \max(v_0, v_1)]$, as required. ■

Remark 6.4 *According to Corollary 6.1 and Corollary 6.3, in the general case, starting with $v \in [\beta e^m(1+q), 1+q)$, once two consecutive unclipped trajectories are realized, all the subsequent trajectories will be unclipped.*

Lemma 6.5 *In the general case, under the main assumption of $e^m\beta < 1$, the trajectory starting with $v(0) = v_0 = \beta(1+q)$, $y(0) = 1$ requires no more than two jumps.*

Proof. Suppose the trajectory is unclipped. Starting with $v_0 = \beta(1+q)$, let us examine the value of $\beta\varphi(\beta(1+q)) = \beta^3(1+q)e^{m(\hat{s}+1)}$, where $\hat{s}$ is the single positive solution to $L(s, \beta, m) = 0$, where $L(s, \beta, m) = \beta(e^{ms} - e^{-s}) + (m+1)(e^{-s} - 1)$ according to (6.34). Hence, $\beta^3(1+q)e^{m(\hat{s}+1)}$ is the value of v_1 after two instant jumps, if starting with $v_0 = \beta(1+q)$. Define the increasing (with respect to s) auxiliary function $C(s) = \beta^3(1+q)e^{m(s+1)}$. One can easily check that the behaviour of $L(s, \beta, m)$ is similar to that of $G_2(s, \beta, m)$ in the proof of Lemma 6.4, in the sense that it decreases firstly from zero and then increases up to infinity, with respect to s.

Let us show that $L(\tilde{s}(\beta, m), \beta, m) > 0$, where $\tilde{s}(\beta, m)$ is the single positive solution to $C(s) = 1+q$: $\tilde{s}(\beta, m) = -\frac{3}{m}\ln\beta - 1$.

Now

$$L(\tilde{s}(\beta, m), \beta, m) = \beta^{-2}e^{-m} - \beta^{\frac{3+m}{m}}e + (m+1)(\beta^{\frac{3}{m}}e - 1).$$

Immediately $L(\tilde{s}(\beta, m), \beta, m) \to \infty$ as $\beta \to 0$. And as $\beta \to e^{-m}$, $L(\tilde{s}(\beta, m), \beta, m) \to e^m - e^{-2-m} + (m+1)(e^{-2} - 1) > 0$, as can be verified easily. Now one can calculate the partial derivative

$$\frac{\partial L(\tilde{s}(\beta,m),\beta,m)}{\partial\beta} = -2\beta^{-3}e^{-m} - \frac{m+3}{m}\beta^{\frac{3}{m}}e + (m+1)\frac{3}{m}\beta^{\frac{3}{m}-1}e.$$

Immediately $\frac{\partial L(\tilde{s}(\beta,m),\beta,m)}{\partial\beta} \to -\infty$ as $\beta \to 0$. Moreover,

$$\lim_{\beta\to e^{-m}}\left(\frac{\partial L(\tilde{s}(\beta,m),\beta,m)}{\partial\beta}\right) = -2e^{2m} - \frac{3+m}{m}e^{-2} + \frac{3(m+1)}{m}e^{m-2} = L_1(m) < 0$$

for any $m > 0$. To see this, by L'Hôpital's rule one can easily check that $L_1(m) \to -2+5e^{-2} < 0$ as $m \to 0$, and obviously $L_1(m) \to -\infty$ as $m \to \infty$. Now calculate the derivative

$$\begin{aligned}
\frac{dL_1(m)}{dm} &= -4e^{2m} - \frac{m-3-m}{m^2}e^{-2} + \frac{3m-3m-3}{m^2}e^{m-2} + \frac{3(m+1)}{m}e^{m-2} \\
&= -4e^{2m} - \frac{-3}{m^2}e^{-2} + \frac{-3}{m^2}e^{m-2} + \frac{3+3m}{m}e^{m-2} \\
&= \frac{1}{m^2}\left\{-4m^2e^{2m} + 3e^{-2} - 3e^{m-2} + (3m^2+3m)e^{m-2}\right\} \\
&= \frac{1}{m^2}L_2(m),
\end{aligned}$$

where $L_2(m) \to 0$ as $m \to 0$, and $L_2(m) \to -\infty$ as $m \to \infty$, and

$$\begin{aligned}
\frac{dL_2(m)}{dm} &= -8me^{2m} - 8m^2e^{2m} - 3e^{m-2} + (6m+3)e^{m-2} + (3m^2+3m)e^{m-2} \\
&= e^{m-2}\left\{-8me^{m+2} - 8m^2e^{m+2} + 9m + 3m^2\right\} \\
&< 0.
\end{aligned}$$

Hence $L_2(m) < 0 \Rightarrow L_1'(m) < 0 \Rightarrow L_1(m) < 0$. That is to say, $\frac{\partial L(\tilde{s}(\beta,m),\beta,m)}{\partial\beta} < 0$ as $\beta \to e^{-m}$.

Let us analyze the second order partial derivative

$$\begin{aligned}
\frac{\partial^2 L(\tilde{s}(\beta,m),\beta,m)}{\partial\beta^2} &= 6\beta^{-4}e^{-m} - \frac{3+m}{m}\frac{3}{m}\beta^{\frac{3}{m}-1}e + \frac{3(m+1)}{m}\frac{3-m}{m}\beta^{\frac{3}{m}-2}e \\
&= \frac{\beta^{\frac{3}{m}-2}}{m^2}e\Big\{6m^2\beta^{-2-\frac{3}{m}}e^{-m-1} - (3+m)3\beta \\
&\quad +3(m+1)(3-m)\Big\} \\
&= \frac{\beta^{\frac{3}{m}-2}}{m^2}eL_3(\beta,m).
\end{aligned}$$

Obviously $L_3(\beta,m) \to \infty$ as $\beta \to 0$. Consider when $\beta \to e^{-m}$

$$\begin{aligned}
\lim_{\beta \to e^{-m}} L_3(\beta,m) &= 6m^2e^{2m+3}e^{-m-1} - (9+3m)e^{-m} + 3(m+1)(3-m) \\
&= 6m^2e^{m+2} - (9+3m)e^{-m} + (3m+3)(3-m) \\
&= 6m^2e^{m+2} - (9+3m)e^{-m} + 9m - 3m^2 + 9 - 3m \\
&= L_4(m),
\end{aligned}$$

where $L_4(m) \to 0$ as $m \to 0$, $L_4(m) \to \infty$ as $m \to \infty$, and

$$\begin{aligned}
\frac{dL_4(m)}{dm} &= 12me^{m+2} + 6m^2e^{m+2} - 3e^{-m} + (9+3m)e^{-m} + 6 - 6m \\
&= (12m + 6m^2)e^{m+2} + (6+3m)e^{-m} + 6 - 6m \\
&> 0.
\end{aligned}$$

Hence, $L_3(\beta,m) > 0$ as $\beta \to e^{-m}$.

Finally one can check that the partial derivative

$$\frac{\partial L_3(\beta,m)}{\partial \beta} = (-2 - \frac{3}{m})6m^2e^{-m+1}\beta^{-3-\frac{3}{m}} - 3(3+m) < 0.$$

Now it only remains to recognize that

$$\begin{aligned}
&\frac{\partial L_3(\beta,m)}{\partial \beta} < 0 \Rightarrow L_3(\beta,m) > 0 \Rightarrow \frac{\partial^2 L(\tilde{s}(\beta,m),\beta,m)}{\partial \beta^2} > 0 \\
\Rightarrow\ &\frac{\partial L(\tilde{s}(\beta,m),\beta,m)}{\partial \beta} < 0 \Rightarrow L(\tilde{s}(\beta,m),\beta,m) > 0,
\end{aligned}$$

as required.

Hence $\hat{s} < \tilde{s}(\beta,m)$ and $C(\hat{s}) < 1+q$ meaning that $\beta\varphi(\beta(1+q)) < 1+q$.

If the trajectory is clipped then the value after the next two instantaneous jumps is even smaller than $\beta\varphi(\beta(1+q))$. ■

Corollary 6.4 *If $e^m\beta < 1$ then no cycle has more than two instantaneous jumps.*

Proof. It is sufficient to notice that, after any instantaneous series of jumps, $v \geq \beta(1+q)$ and $\beta\varphi(v) \leq \beta\varphi(\beta(1+q))$. ■

Lemma 6.6 *In the unconstrained case, 2−cycles are absent iff $\beta < \bar{\beta}$, where $\bar{\beta}$ is the single solution in the interval $(0, e^{-m})$ to equation (6.7).*

Proof. Clearly a 2−cycle, described by the starting point

$$V_0^{(2)} = \frac{(1+q)e^m\beta^2\left(1 - e\beta^{\frac{2}{m}}\right)(m+1)}{1 - e^{m+1}\beta^{(2+\frac{2}{m})}},$$

does not exist iff $\frac{V_0^{(2)}}{\beta} < 1+q$. (Compare with the proof of Lemma 6.1.) Or equivalently,

$$Q_1(\beta) = (m+1)e^m\beta(1-e\beta^{\frac{2}{m}}) + e^{m+1}\beta^{2(1+\frac{1}{m})} < 1.$$

Let us check the behaviour of $Q_1(\beta)$. Immediately, $Q_1(\beta,m) \to 0$ as $\beta \to 0$, and $Q_1(\beta,m) \to 1$ as $\beta \to e^{-\frac{m}{2}}$. And one can analyze the partial derivative

$$\begin{aligned}
\frac{\partial Q_1(\beta,m)}{\partial\beta} &= (m+1)\left\{e^m(1-e\beta^{\frac{2}{m}}) - e^m\beta e\frac{2}{m}\beta^{\frac{2}{m}-1}\right\} \\
&\quad + e^{m+1}\beta^{1+\frac{2}{m}}2(1+\frac{1}{m}) \\
&= e^m\left\{(m+1)\left\{(1-e\beta^{\frac{2}{m}}) - e\beta^{\frac{2}{m}}\frac{2}{m}\right\}\right. \\
&\quad \left. + e\beta^{\frac{2}{m}}\beta 2(1+\frac{1}{m})\right\} \\
&= e^m Q_2(\beta,m).
\end{aligned}$$

It is noted that $Q_2(\beta,m) \to (m+1)(k-1) > 0$ as $\beta \to 0$, and $Q_2(\beta,m) \to (e^{-\frac{m}{2}} - 1)(\frac{2}{m}+2) < 0$ as $\beta \to e^{-\frac{m}{2}}$, indicating that $\frac{\partial Q_1(\beta,m)}{\partial\beta} \to 0$ as $\beta \to 0$, and $\frac{\partial Q_1(\beta,m)}{\partial\beta} < 0$ as $\beta \to e^{-\frac{m}{2}}$. Furthermore, one can analyze now the partial derivative

$$\begin{aligned}
\frac{\partial Q_2(\beta,m)}{\partial\beta} &= (m+1)\left\{-e\frac{2}{m}\beta^{\frac{2}{m}-1} - e(\frac{2}{m})^2\beta^{\frac{2}{m}-1}\right\} + \\
&\quad (\frac{2}{m}+1)(1+\frac{1}{m})2e\beta^{\frac{2}{m}} \\
&= \beta^{\frac{2}{m}-1}e\left\{(m+1)\left\{-\frac{2}{m}-(\frac{2}{m})^2\right\} + (\frac{2}{m}+1)(1+\frac{1}{m})2\beta\right\} \\
&= \beta^{\frac{2}{m}-1}e\left\{\frac{-m(m+1)2-4(m+1)+(2+m)(m+1)2\beta}{m^2}\right\} \\
&= \frac{\beta^{\frac{2}{m}-1}e(m+1)2}{m^2}[-m-2+(2+m)\beta] \\
&= \frac{\beta^{\frac{2}{m}-1}e(m+1)2}{m^2}[m(-1+\beta)+2(\beta-1)] < 0,
\end{aligned}$$

where the inequality holds as $\beta < e^{-\frac{m}{2}} < 1$.

The above analysis of derivatives shows that $Q_1(\beta)$ increases with β from $Q_1(0) = 0$ and, after a single stationary point, decreases up to $Q_1(e^{-\frac{m}{2}}) = 1$. Therefore, equation (6.7) has a single solution in interval $(0, e^{-\frac{m}{2}}) \supset (0, e^{-m})$.

One can easily check that

$$\left.\frac{V_0^{(2)}}{\beta}\right|_{\beta=e^{-m}} > 1+q, \tag{6.40}$$

which is equivalent to $\bar{\beta} < e^{-m}$. ■

Remark 6.5 *Suppose $\beta < e^{-m}$. Then, in the general case, $2-$cycles exist for some (big enough) values of B iff $\beta \geq \bar{\beta}$. According to Lemma 6.1, $1-$cycles also exist. What is actually realized, depends on the initial conditions $v(0) = v_0, y(0) = 1$.*

Lemma 6.7 *In the unconstrained case, the continuous trajectory starting from $v(0) = v_0 = \beta(1+q), y(0) = 1$ reaches level $y(\hat{s}) = 1$ with such a value of $v(\hat{s})$ that $\beta v(\hat{s}+1) < 1+q$ if and only if $\beta < \bar{\beta}$.*

Proof. Clearly $\beta v(\hat{s}+1) = \varphi(v_0) = \beta^2(1+q)e^{m(\hat{s}+1)}$, where $\hat{s} > 0$ solves equation (6.34) at $v = v_0 = \beta(1+q)$. Now $\beta v(\hat{s}+1) < 1+q \Leftrightarrow \beta^2 e^{m(\hat{s}+1)} < 1$.

Firstly, one can check that equations

$$\beta^2 e^{m(\hat{s}+1)} = 1; \tag{6.41}$$

$$\beta(e^{m\hat{s}} - e^{-\hat{s}}) + (m+1)(e^{-\hat{s}} - 1) = 0 \tag{6.42}$$

hold iff $\beta = \bar{\beta}$. Indeed, substitute expression $\hat{s} = \frac{-2\ln\beta}{m} - 1$ obtained from (6.41), into (6.42):

$$\beta(\frac{e^{-m}}{\beta^2} - e\beta^{2/m}) + (m+1)(e\beta^{2/m} - 1) = 0 \Leftrightarrow (6.7) \Leftrightarrow \beta = \bar{\beta}.$$

Secondly, from (6.42) we obtain

$$\frac{d\hat{s}}{d\beta} = \frac{e^{m\hat{s}} - e^{-\hat{s}}}{(m+1)e^{-\hat{s}} - \beta(me^{m\hat{s}} + e^{-\hat{s}})}.$$

The numerator is positive for $\hat{s} > 0$. After we substitute $\beta = \frac{(1-e^{-\hat{s}})(m+1)}{e^{m\hat{s}} - e^{-\hat{s}}}$, obtained from (6.42), into the denominator, we have

$$\frac{(m+1)e^{-\hat{s}}(e^{m\hat{s}} - e^{-\hat{s}}) - (1-e^{-\hat{s}})(m+1)(me^{m\hat{s}} + e^{-\hat{s}})}{e^{m\hat{s}} - e^{-\hat{s}}} < 0$$

because $e^{m\hat{s}-\hat{s}} - me^{m\hat{s}} + me^{m\hat{s}-\hat{s}} - e^{-\hat{s}} < 0$ at any positive $\hat{s}$ and m. More exactly,

$$\lim_{m\to 0}(e^{m\hat{s}-\hat{s}} - me^{m\hat{s}} + me^{m\hat{s}-\hat{s}} - e^{-\hat{s}}) = 0,$$

$$\frac{\partial}{\partial m}(e^{m\hat{s}-\hat{s}} - me^{m\hat{s}} + me^{m\hat{s}-\hat{s}} - e^{-\hat{s}}) = e^{m\hat{s}}(\hat{s}e^{-\hat{s}} - m\hat{s} - 1 + e^{-s}(1+m\hat{s})) = e^{m\hat{s}}M(m,\hat{s}),$$

$$\frac{\partial M(m,\hat{s})}{\partial m} = -\hat{s} + e^{-\hat{s}}\hat{s} < 0.$$

Thus $\frac{d\hat{s}}{d\beta} < 0$.

Finally, we intend to prove that $\lim_{\beta\to 0}\beta^2 e^{m(\hat{s}+1)} = 0$. When $\beta \to 0$, $\hat{s}$ increases, but the limit cannot be finite. (Otherwise, passing to the limit in (6.42) would imply $(m+1)(e^{-\lim_{\beta\to 0}\hat{s}} - 1) = 0$.) Hence $\lim_{\beta\to 0}\hat{s} = \infty$, and from (6.42) we have $\lim_{\beta\to 0}\beta e^{m\hat{s}} = m+1 \Rightarrow \lim_{\beta\to 0}\beta^2 e^{m(\hat{s}+1)} = 0$. Therefore, $\beta^2 e^{m(\hat{s}+1)} < 1 \Leftrightarrow \beta < \bar{\beta}$ because $\bar{\beta}$ is the single value of β providing $\beta^2 e^{m(\hat{s}+1)} = 1$, and the left hand side is obviously a continuous function of β. ■

Proof of Theorem 6.1. Note that $\bar{\beta}$ is the single solution to equation (6.7) in the interval $(0, e^{-m})$ according to Lemma 6.6. If $\beta < e^{-m}$ Lemma 6.5 excludes trajectories with three or more instantaneous jumps (perhaps after one first continuous trajectory is realised).

(a) Suppose $\beta < \bar{\beta}$ and $B \geq B^* \Leftrightarrow q \leq q^*$. Lemma 6.7 implies that (perhaps after the first one instantaneous series of jumps) multiple reductions of component v never occur and all the further values of v_i belong to $[\beta e^m(1+q), 1+q)$. For the proof of the latter statement, remember the denotations introduced before Lemma 6.4, and equality $\varphi(1+q) = \beta e^m(1+q)$. Even if a continuous trajectory starting from $v(0) = v_i, y(0) = 1$ is clipped, $v_{i+1} = \varphi(\hat{v}_0) \in [\beta e^m(1+q), 1+q)$, where $\hat{v}_0$ was defined in Corollary 6.1.

Suppose there exists a clipped continuous trajectory starting from $v(0) = v_i, y(0) = 1$. (Actually, i can equal 1 or 2.) The next trajectory starting from $v(0) = v_{i+1} \in [\beta e^m(1+q), 1+q), y(0) = 1$ cannot be clipped because otherwise we would have obtained a clipped $1-$cycle which contradicts the last statement of Lemma 6.2. Thus $v_{i+1} \geq \hat{v}_0$ and $v_{i+2} = \varphi(v_{i+1})$. Since $v_{i+1} = \varphi(\hat{v}_0)$, we can use Corollary 6.3: $v_{i+2} \geq \hat{v}_0$, so that trajectory starting from $v(0) = v_{i+2}, y(0) = 1$ is also unclipped. According to Remark 6.4, all the subsequent trajectories are unclipped and converge to the limiting unclipped $1-$cycle in accordance with Lemma 6.4.

If $\beta \geq \bar{\beta}$ then, according to Remark 6.5, statement (a) is false.

All the presented reasoning holds also if $\bar{\beta} \leq \beta < e^{-m}$ and $v_0 \in [\beta e^m(1+q), 1+q)$: multiple jumps never occur and $\forall i \geq 0\ v_i \in [\beta e^m(1+q), 1+q)$. (See the proof of Lemma 6.4.) On the other hand, according to Remark 6.5, for some initial conditions, a simple $2-$cycle can be realized if B is big enough. This observation justifies Remark 6.1.

(b) If $B = B^*$, the previous paragraph is correct, but (independently of the initial state) no-one continuous trajectory can have multiple jumps at the end, because it cannot be situated below the curve starting from $v(0) = q, y(0) = 0$ which results in the single jump at the end. Thus, trajectories converge to the $1-$cycle that is critical according to Lemma 6.2. The necessity of inequality $\beta < e^{-m}$ can be proved similarly to the part (c).

(c) Similarly to case (b), continuous trajectories having multiple jumps at the end cannot be realized if $B < B^* \Leftrightarrow q > q^*$. According to Lemma 6.2, one cannot meet an unclipped $1-$cycle. Corollary 6.1 and Lemma 6.4 imply that $\hat{v}_0 \in [\beta e^m(1+q), 1+q)$. Moreover, $\varphi(\hat{v}_0) < \hat{v}_0$ because otherwise, starting from $v(0) = \hat{v}_0, y(0) = 1$ we would have had two consecutive unclipped trajectories leading to an unclipped limiting $1-$cycle according to Remark 6.4 and Lemma 6.4.

Now one of the following two scenarios can take place.

If $v_0 < \hat{v}_0$, then the first continuous trajectory is clipped and $v_1 = \varphi(\hat{v}_0) < \hat{v}_0$, so that the next continuous trajectory is also clipped, and the limiting clipped $1-$cycle is attained after one iteration.

If $v_0 \geq \hat{v}_0$, then the first continuous trajectory is unclipped, but $v_1 = \varphi(v_0) < \hat{v}_0$. (Otherwise we face two consecutive unclipped trajectories leading to the existence of an unclipped $1-$cycle.) Hence v_1 gives a clipped continuous trajectory, and, according to the previous paragraph, we finish with the clipped $1-$cycle attained after two iterations.

As Lemma 6.1 says, an unclipped $1-$cycle does not exist if $\beta e^m < 1$. One can easily show that inequality $\beta e^m < 1$ is also necessary for the existence of clipped $1-$cycles. Indeed, if $\beta e^m \geq 1$ then formula (6.28) gives $S \leq 0$, and that formula remains the same for clipped and unclipped cycles because equation (6.26) is universal.

The very last statement is justified in full by all the previous reasoning. ■

Before proving Proposition 6.1, let us justify formula (6.9). Clearly,

$$T_{cycle} = \int_0^{T_{cycle}} dt = \int_0^{S+1} \left\{ T + \frac{By(s)}{\mu} \right\} ds,$$

where S is given by (6.28), and expression (6.9) now follows from the calculation

$$T_{cycle} \;\; = \;\; TS + \frac{B}{\mu}\int_0^S y(s)ds + T + \frac{B}{\mu}$$

$$
\begin{aligned}
&= \frac{T}{q}\int_0^S \left\{ \frac{V_0}{m+1}e^{ms} - q + e^{-s}\left(1+q-\frac{V_0}{m+1}\right)\right\} ds + \frac{T}{q} + \frac{T}{m}\ln\frac{1}{\beta} \\
&= \frac{T}{m}\ln\frac{1}{\beta} + \frac{T}{q}\left\{ 1 + \frac{(1+q)(1-e\beta^{\frac{1}{m}})}{m(1-e^{m+1}\beta^{1+\frac{1}{m}})} - \frac{q}{m}\ln\frac{1}{\beta} + q \right. \\
&\quad - e\beta^{\frac{1}{m}}\left(1+q-\frac{(1+q)e^{m}\beta(1-e\beta^{\frac{1}{m}})}{1-e^{m+1}\beta^{1+\frac{1}{m}}}\right) - \frac{(q+1)e^{m}\beta(1-e\beta^{\frac{1}{m}})}{m(1-e^{m+1}\beta^{1+\frac{1}{m}})} \\
&\quad \left. + \left(1+q-\frac{(1+q)e^{m}\beta(1-e\beta^{\frac{1}{m}})}{1-e^{m+1}\beta^{1+\frac{1}{m}}}\right)\right\} \\
&= \frac{T(1+q)}{qm}\left\{ m + \frac{1-e\beta^{\frac{1}{m}}}{1-e^{m+1}\beta^{1+\frac{1}{m}}} \right. \\
&\quad - me\beta^{\frac{1}{m}}\left(1-\frac{e^{m}\beta(1-e\beta^{\frac{1}{m}})}{1-e^{m+1}\beta^{1+\frac{1}{m}}}\right) - \frac{e^{m}\beta(1-e\beta^{\frac{1}{m}})}{1-e^{m+1}\beta^{1+\frac{1}{m}}} \\
&\quad \left. + m - \frac{me^{m}\beta(1-e\beta^{\frac{1}{m}})}{1-e^{m+1}\beta^{1+\frac{1}{m}}}\right\} \\
&= \frac{T(1+q)}{qm}\left\{ m + \frac{m+1-e\beta^{\frac{1}{m}} - me\beta^{\frac{1}{m}} + me^{m+1}\beta^{1+\frac{1}{m}} - e^{m}\beta}{1-e^{m+1}\beta^{1+\frac{1}{m}}} \right. \\
&\quad \left. + \frac{e^{m+1}\beta^{1+\frac{1}{m}} - me^{m}\beta}{1-e^{m+1}\beta^{1+\frac{1}{m}}}\right\} \\
&= \frac{T(1+q)}{qm}\left\{ m + \frac{(1-e\beta^{\frac{1}{m}})(m+1)(1-e^{m}\beta)}{1-e^{m+1}\beta^{1+\frac{1}{m}}}\right\} \\
&= \frac{T(B+\mu T)}{\mu T m}\left\{ m + \frac{(1-e\beta^{\frac{1}{m}})(m+1)(1-e^{m}\beta)}{1-e^{m+1}\beta^{1+\frac{1}{m}}}\right\}.
\end{aligned}
$$

Proof of Proposition 6.1. In the case $q \le q^* \Leftrightarrow B \ge B^* = \frac{\mu T}{q^*}$, the cycle is unclipped. The average sending rate can be calculated according to formula $\bar{\lambda} = \frac{\int_0^{S+1} w(s)ds}{T_{cycle}} = \frac{B}{T_{cycle}}\int_0^{S+1} V_0 e^{ms} ds$ (see (6.24)). The average throughput can be calculated as the following.

$$
\begin{aligned}
\bar{g} &= \frac{1}{T_{cycle}}\left\{\int_0^{T_{cycle}-T-\frac{B}{\mu}} \lambda(t)dt + \mu\left(T+\frac{B}{\mu}\right)\right\} \\
&= \frac{1}{T_{cycle}}\left\{\int_0^{S} w(s)ds + \mu\left(T+\frac{B}{\mu}\right)\right\} \\
&= \frac{1}{T_{cycle}}\left\{B\int_0^{S} v(s)ds + \mu T + B\right\}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{T_{cycle}}\left\{B\int_0^S V_0 e^{ms}ds+\mu T+B\right\} \\
&= \frac{\mu T}{qmT_{cycle}}\left\{V_0(\frac{1}{e^m\beta}-1)+qm+m\right\} \\
&= \frac{\mu T(q+1)}{qmT_{cycle}}\left\{m+\frac{(1-e\beta^{\frac{1}{m}})(m+1)(1-e^m\beta)}{1-e^{m+1}\beta^{1+\frac{1}{m}}}\right\}=\mu.
\end{aligned}
$$

Also, the average amount of data in the buffer is calculated as below.

$$
\begin{aligned}
\bar{x} &= \frac{1}{T_{cycle}}\int_0^{T_{cycle}} x(t)dt = \frac{1}{T_{cycle}}\int_0^{S+1} x(s)(T+\frac{x(s)}{\mu})ds \\
&= \frac{1}{T_{cycle}}\left\{\int_0^S x(s)(T+\frac{x(s)}{\mu})ds+B(T+\frac{B}{\mu})\right\} \\
&= \frac{1}{T_{cycle}}\left\{TB\int_0^S y(s)ds+\frac{B^2}{\mu}\int_0^S y^2(s)ds+BT+\frac{B^2}{\mu}\right\}.
\end{aligned}
$$

■

In case $B < B^* \Leftrightarrow q > q^*$, the cycle is clipped. As before, we use $t(s)$ for the original (new) time scale. The graph of the cycle in the plane (v,y) looks similarly to Figure 6.1; one has only to replace "Buffer size (B)" on the $y-$axis by 1. Suppose the cycle starts at $s = S_A = 0$ from point A, reaches point B at time moment S_B and so on. We shall use denotations like S_{BC} for $S_C - S_B$.

Point C has coordinates $y = 0$ and $v = q$, so that, when $s \in [S_C, S_D]$,

$$y(s) = \frac{q}{m+1}e^{m(s-S_C)} - q + e^{-(s-S_C)}(\frac{mq}{m+1})$$

according to (6.25). Therefore, $S_{CD} = S_D - S_C$ is the single positive solution to equation $A(s) = 0$ where

$$A(s) = qe^{ms} + e^{-s}qm - (m+1)(q+1).$$

(Note that $\lim_{s\to 0} A(s) = -1-m, \lim_{s\to\infty} A(s) = \infty$ and $\frac{dA}{ds} > 0$.) Equation (6.13) is proved. Formula $V_0^{clipped} = v(0) = \frac{\mu T}{B}\beta e^{m(S_{CD}+1)}$ at the beginning of the cycle follows from (6.24), so that expression (6.12) is justified. According to (6.25),

$$y(s) = \frac{V_0^{clipped}}{m+1}e^{ms} - q + e^{-s}\left(1+q-\frac{V_0^{clipped}}{m+1}\right)$$

for $s \in [0, S_B]$, where $S_B = S_{AB}$ is the minimal positive solution of equation $y_{AB}(S_{AB}) = 0$. (The maximal solution is phantom, corresponding to the last moment when component y equals zero in case we ignore the non-negativity constraint, i.e., if we deal with the unconstrained case.) Equation (6.15) is obtained.

Now

$$T_{cycle} = \int_0^{S+1} \left[T + \frac{By(s)}{\mu} \right] ds = \frac{T}{m} \ln \frac{1}{\beta} + \frac{B}{\mu} \left\{ \int_0^{S_B} y(s) ds + \int_{S_C}^{S_D} y(s) ds + 1 \right\}$$

and formulae (6.16) (6.17) (6.14) are proved where we made the trivial change of the (new) time scale:

$$Y_{CD}(s) = y(S_C + s); Y_{AB}(s) = y(S_A + s) = y(s).$$

Proof of Proposition 6.2. Similarly to the case $B \geq B^*$, $\bar{\lambda} = \frac{1}{T_{cycle}} \int_0^{S+1} w(s) ds$ leads to formula (6.18).

For the average throughput we have, using (6.24):

$$\begin{aligned}
\bar{g} &= \frac{1}{T_{cycle}} \left\{ \int_0^{S_{AD}} \lambda(s) \left(T + \frac{By(s)}{\mu} \right) ds + \mu \left(T + \frac{B}{\mu} \right) \right\} \\
&= \frac{1}{T_{cycle}} \left\{ \int_0^{S_{AD}} w(s) ds + \mu (T + \frac{B}{\mu}) \right\} \\
&= \frac{1}{T_{cycle}} \left\{ B \int_0^{S} v(s) ds + \mu T + B \right\} \\
&= \frac{1}{T_{cycle}} \left\{ B \left(\frac{V_0^{clipped}}{m} e^{mS} - \frac{V_0^{clipped}}{m} \right) + \mu T + B \right\} \\
&= \frac{1}{T_{cycle}} \left\{ \frac{BV_0^{clipped}}{m} (\frac{1}{\beta e^m} - 1) + \mu T + B \right\}.
\end{aligned}$$

Also the average amount of data in the buffer is calculated as follows.

$$\begin{aligned}
\bar{x} &= \frac{1}{T_{cycle}} \int_0^{T_{cycle}} x(t) dt \\
&= \frac{1}{T_{cycle}} \left\{ \int_0^{S_B} By(s) \left(T + \frac{By(s)}{\mu} \right) ds + \int_{S_C}^{S_D} By(s) \left(T + \frac{By(s)}{\mu} \right) ds \right. \\
&\quad \left. + B \left(T + \frac{B}{\mu} \right) \right\} \\
&= \frac{1}{T_{cycle}} \left\{ TB \left(\int_0^{S_{AB}} Y_{AB}(s) ds + \int_0^{S_{CD}} Y_{CD}(s) ds \right) \right. \\
&\quad \left. + \frac{B^2}{\mu} \left(\int_0^{S_{AB}} Y_{AB}^2(s) ds + \int_0^{S_{CD}} Y_{CD}^2(s) ds \right) + B \left(T + \frac{B}{\mu} \right) \right\}.
\end{aligned}$$

■

Chapter 7

General conclusion

In this chapter, let us summarize this thesis as follows. In a nutshell, this thesis revolves about the three problem sets raised in Chapter 1. In this way, we firstly tackled the solvability issue of CTMDPs in Chapter 2 (see Problem set (a) in Chapter 1). There, for a discounted model in a Borel state space, when there is only one performance criterion (that is, in the unconstrained case), we firstly developed the dynamic programming approach by showing the existence and uniqueness (up to the class of w'-bounded functions, where w' is a bounded function) of a solution to the Bellman equation, studied the relations between the Bellman equation and the DLP, and established the existence of a deterministic stationary optimal policy. When there are more than one criteria (that is, in the constrained case), the dynamic programming approach often turns inconvenient, and thus we developed the convex analytic approach, which studies the original CTMDP optimization problem via an equivalent problem in the form of a PLP in the space of special measures, known as occupation measures. For this PLP, firstly, we studied its relations with its DLP by showing the absence of duality gap, which in some sense, allows one to conclude the duality between the convex analytic approach and the dynamic programming approach; and secondly, with the w'-weak topology, we showed the compactness of the feasible region as well as the lower semicontinuity of the objective function with respect to the occupation measure, which then finally led to the existence of a (possibly randomized) stationary optimal policy. The conditions assumed here were standard, but did allow unbounded transition rates and cost rates, as illustrated by an example. It should be noted that the solvability of CTMDPs has never been addressed by the moment of writing up this thesis, as soon as the features of a Borel state space, unbounded transition rates and cost rates, an arbitrary finite number of constraints, and general history-dependent policies meet together.

Secondly we responded to Problem set (b) by addressing the approximation issues of CTMDPs. Indeed, Chapter 3, Chapter 4 and Chapter 5 are all on the fluid approximations of CTMDPs, which allow one to obtain a policy for the underlying CTMDP with a satisfactory performance (that is, an AFO policy) by solving its appropriately formulated fluid model. Such fluid approximations have been widely used in engineering problems, and in the context of optimization, they are often verified by simulations. However, in the sense of fluid scaling, the difference between the performance functional in the stochastic model and that in the fluid model can be estimated with explicit formulae involving only primitives, as done in this thesis. To our best knowledge, while the comparison of the trajectories in the stochastic model with those in the fluid model has been rich in the literature, that at the level of performance functionals has only received relatively limited attention, especially when the estimate of the difference is concerned. Nevertheless, this difference can be indeed important and useful, as it can be taken as a measure when comparing different AFO policies. Indeed, this difference is often read as the efficiency of the underlying AFO policy.

For this, in Chapter 3 for a CTMDP with total expected cost and an absorbing state, which can be read as a one-dimensional Birth-and-Death process or a controlled $M/M/1$ queueing system, we firstly proposed an appropriate fluid model, which can differ from the naturally looking one, as confirmed by means of an example. Then, for this fluid model we showed a feedback (state-dependent) translation of the fluid optimal policy results in an AFO and AO policy, with its efficiency being estimated explicitly involving only the primitives. While they are interesting in their own right, one direct application of this type of results obtained in Chapter 3 can be found in a special CTMDP with a long run expected average, as illustrated by Chapter 4, which studies the same issues of fluid approximations of inventory level-dependent EOQ and EPQ models. In this link, it is not too harmful or disgraceful to regard Chapter 4 as a big corollary of Chapter 3. The translations in Chapter 3 and Chapter 4 are both of feedback type, and their efficiencies are both $O(\frac{1}{n})$, where n is the fluid scaling parameter. An alternative one is the tracking policy, as considered in Chapter 5. The tracking policy is resulted in by a time-dependent but nearly state-independent translation of the fluid optimal policy. For a special finite CTMDP in the form of a bandwidth sharing network for which one aims at the optimal resource application to minimize the expected total holding cost, we studied the efficiency of the tracking policy. By showing the existence of a fluid optimal policy, which is in the form of a piecewise constant function in time, we firstly dealt with a general bandwidth-sharing network, where by general is meant no extra conditions being imposed on the network parameters, for which we showed the efficiency of the tracking policy to be $O(\frac{1}{\sqrt{n}})$; and secondly, after imposing some extra conditions on the network parameters so that the time horizon is relatively small, we showed that the tracking policy could be also of the efficiency $O(\frac{1}{n})$. By means of an example, we are assured the feedback translation cannot generally result in an AFO policy of the efficiency better than $O(\frac{1}{n})$. Therefore, the observation of a situation leading to the efficiency of $O(\frac{1}{n})$ for the tracking policy is indeed putting the tracking policy in a favoured position when compared with the feedback translation, simply because of the less information required to implement it. As a matter of fact, the efficiency of the tracking policy was only mentioned briefly in Gajrat et al (2003) to be $O(\frac{1}{\sqrt{n}})$ in a different problem (a tandem of two queues in the discrete time), and has been thus regarded preliminarily less efficient than the feedback translations. Therefore, the observation of the efficiency of $O(\frac{1}{n})$ for the tracking policy is deemed interesting.

One should feel free to regard Chapter 3, Chapter 4 and Chapter 5 as formal justifications of fluid approximations in their specific settings. Given such preliminary beliefs in fluid approximations, in Chapter 6 we investigated a fluid model of an Internet router under the MIMD congestion control, specifically corresponding to STCP. This is a response to Problem set (c). In greater detail, by formulating a fluid model taking into account the dynamics of congestion window, consisting of a continuous part and a discrete part in the form of an instantaneous reduction (that is, a hybrid model), we firstly studied the long run behaviour of trajectories (the number of packets in the router along the time), for which a limiting regime, in the form of a cycle, was shown. This limiting regime then allowed one to compute the performance criteria (dependent on the buffer size) in the forms of two long run averages by focusing on only one cycle. We then studied the optimal buffer size, given as a Pareto set, as usual when dealing with problems with multi-objectives. The most interesting case of this model is when the parameters are tuned to correspond to STCP, for which simulation results (from Dr. Urtzi Ayesta) can confirm the obtained results and verify the underlying fluid model. Indeed, even though the buffer sizing problem (see problem set (c) also) has been intensively studied for TCP New Reno, for STCP it was not much addressed. One such example is Ayesta et al (2008), where at an early stage of this research we considered a simpler and rougher fluid model [1]. Since the current TCP New Reno is under various challenges mainly for its incapability of using the high speed network capacity, it makes great sense to consider fundamental problems such as the buffer sizing problem in the network design for its possible alternatives such as STCP, which better fits high speed networks, and such an attempt was carried out in Chapter 6.

As for the future, the following extensions and continuations of this thesis are of particular interest to us. Firstly, we believe that the approaches developed in Chapter 2 can be applied to CTMDPs in other settings, such as the case of long run expected total costs, the case of an absorbing state as well as the case of a finite horizon, though for the latter case, we foresee that the optimal policy is generally not stationary. To start with, one might consider the simpler case of a countable state space. Such extensions will fill in the gaps between Chapter 2 and Chapter 3, Chapter 4 and Chapter 5, respectively, because in the former two (resp. latter one), we indeed restricted ourselves to the class of deterministic stationary policies (resp. deterministic Markov policies). Secondly, given that studies of fluid approximations in Chapter 3, Chapter

[1] For this reason and completeness, we include that model in Appendices.

4 are about problems in one dimensional case, and Chapter 5 was indeed reduced to a one dimensional case, it would be interesting to study similar problems in the multidimensional case. For example, in our opinion, there is little doubt that the approach (via the application of Dynkin's formula) demonstrated in Chapter 3 can be readily applied to justifications of fluid (deterministic) SIR (Susceptible-Infected-Recovered) models in epidemics such as Clancy and Piunovskiy (2005); Piunovskiy and Clancy (2008); Gleissner (1988); Kermack and McKendrick (1927), where one is particularly interested in the process up to the absorption at state zero. The similar approach may also be applied to study the fluid approximations to the scheduling of special single-server-multiclass-jobs networks, where the general optimal solution is already obtained for the fluid model. Thirdly, regarding the EPQ model considered in Chapter 4, it was assumed there that the machine is perfectly reliable. A more realistic assumption would allow the production rate to be dependent on the state of the machine, too. For example, in case the machine is only associated with two states, "on" and "down", if the (random) life time of the machine is sufficiently long relative to the production time, then fluid approximations are often used so that the inventory-level process during the life time is taken to be deterministic, see Berman et al (2007); Chen and Mandelbaum (1994) for instance. A more general case can be that the state of machine itself follows a Markov process with a richer state space. It would be interesting to justify such Markov-modulated fluid models. One example fitting in this line can be found in Altman et al (2001), but again the interesting object for us is the estimate of accuracy. Fourthly, after all, in this thesis the fluid approximations were justified in the sense of fluid scaling. Recently, another direction in this topic is towards justifying fluid approximations not in the sense of fluid scaling, but with respect to the certain values of parameters: for its concerned problem, the fluid approximation was shown to be accurate in Meyn (2005) when there is heavy traffic loading. Undoubtedly, it would be interesting to justify fluid approximations in this sense to other problems. Fifthly, apart from fluid approximations, there are other possibilities to obtain a nearly optimal policy as well as the value of the concerned CTMDP. Particularly, having in hand the PLP for the constrained CTMDP, besides its qualitative properties, it would be interesting to study possible numerical procedures to solve it. On this topic, to our best knowledge, in such an infinite dimensional case, there has been not so many results as in the finite dimensional case. Sixthly, leaving the PLP alone, value iterations and policy iterations have been well known numerical techniques for CTMDPs. Such procedures are subject to potential technical difficulties when the underlying CTMDP is too big, say when the state space and action space are too rich. In particular, to implement the policy iteration, at each step, one needs calculate the performance functional of the underlying model under a fixed policy, and this step has been one of the main sources of difficulties. Recently, for some optimization problems in queueing networks, there have been heuristic results such as Li and Glazebrook (2010) on the situation when

one replaces the performance functional for the underlying stochastic model with that of the corresponding fluid model when carrying out policy iterations. Another similar work is Meyn (1997). We deem it interesting to study more carefully the impact of the fluid approximations in such numerical procedures. Finally, one may consider the formal justification of the fluid model considered in Chapter 6.

Appendices

A rougher fluid model of an Internet router

In this section, we shall consider the same problem as in Chapter 6, but by taking a simpler and rougher model. The results to be presented are based on our early research Ayesta et al (2008). To avoid unnecessary duplications, below we will only sketch the mathematical model and some main results. Accordingly, we excuse ourselves for skipping the sections of introduction and conclusion.

Mathematical model

Let one long-lived MIMD TCP connection send data through a bottleneck router with buffer size B and transmission capacity μ. Denote by $\lambda(t)$ the instantaneous sending rate of connection at time $t \in [0,\infty)$. We consider a fluid model. Namely, data is represented by a fluid that flows into the buffer with rate $\lambda(t)$, and it leaves the buffer with constant rate μ, if there is a backlog in the buffer. Denote by $x(t)$ the amount of data in the buffer at time $t \in [0,\infty)$. Then, the evolution of x is described by the following differential equation

$$\frac{dx(t)}{dt} = \begin{cases} \lambda(t) - \mu, & \text{if } x(t) > 0, \text{or if } x(t) = 0 \text{ and } \lambda(t) > \mu, \\ 0, & \text{if } x(t) = 0 \text{ and } \lambda(t) \le \mu. \end{cases} \tag{1}$$

If $x(t) < B$, the sending rate of the connection increases exponentially in time with rate m. Thus, if $x(t) < B$,

$$\frac{\lambda(t)}{dt} = m\lambda(t). \tag{2}$$

When $x(t)$ reaches B, a congestion signal is sent to one or several TCP connections. Upon the reception of the congestion signal at time t, the TCP connection reduces its sending rate by a multiplicative factor $\beta \in (0,1)$, that is, $\lambda(t+0) = \beta\lambda(t-0)$, where t is a moment of congestion. We call such moments 'jump moments' (of component λ).

Let us now formulate performance criteria. On one hand, we are interested in obtaining as large throughput as possible. That is, we are interested to maximize the average sending rate

$$\bar{\lambda} = \lim_{t\to\infty} \frac{1}{t}\int_0^t \lambda(s)ds.$$

On the other hand, we are interested to make the delay of data in the buffer as small as possible. That is, we are also interested in minimizing the average amount of data in the buffer

$$\bar{x} = \lim_{t\to\infty} \frac{1}{t}\int_0^t x(s)ds.$$

As usual, the solution will be given as a Pareto set, in the form of a trade-off curve on the plane $(\bar{\lambda},\bar{x})$.

For the sake of simplicity we put $m = 1$, $\mu = 1$. In the general case, one should make change of variables: $\tilde{t} = mt$, $\tilde{x} = (m/\mu)x$, $\tilde{\lambda} = \lambda/\mu$. Clearly, $\bar{\lambda}$ and $\bar{x}$ must be multiplied by μ and by μ/m, respectively.

In what follows, we investigate the trajectories of dynamical system (1), (2) which turn to converge to (stable) cycles. Remember that $\lambda(t+0) = \beta\lambda(t-0)$, where t is such that $x(t-0) = B$, $\lambda(t-0) > 1$, and in principle there can be several instant jumps meaning that β above should be replaced by β^k.

Definition .1 *A trajectory of (1), (2) on a finite interval $t \in [0,T]$ is called a cycle if $x(0) = B$, $\lambda(0) < 1$, $x(T) = B$, $\lambda(T) \geq 1$ and $\lambda(0) = \beta^k\lambda(T)$, where $k \geq 1$ is such that $\beta^{k-1}\lambda(T) > 1$. A cycle is called simple if $k = 1$ and $\forall t \in (0,T)$ $x(t) < B$. (One cannot exclude in advance the situation when a cycle has several different loops resulting from the jumps of λ to several different points.) Cycles with component x being zero during a positive time interval are called clipped (Figure 1).*

Actually, a cycle is a one period of the vector-valued periodic function $(x(t), \lambda(t))$, but it is better to represent it graphically as a phase portrait: $x(t)$ against $\lambda(t)$. For clipped cycles, $\bar{\lambda}$ and $\bar{x}$ both increase as B increases. The trade-off of $\bar{x}$ versus $\bar{\lambda}$ is described in Theorem .3.

As it will be shown in Theorem .2, only a simple cycle can exist which is clipped or unclipped depending on the values of parameters B and β.

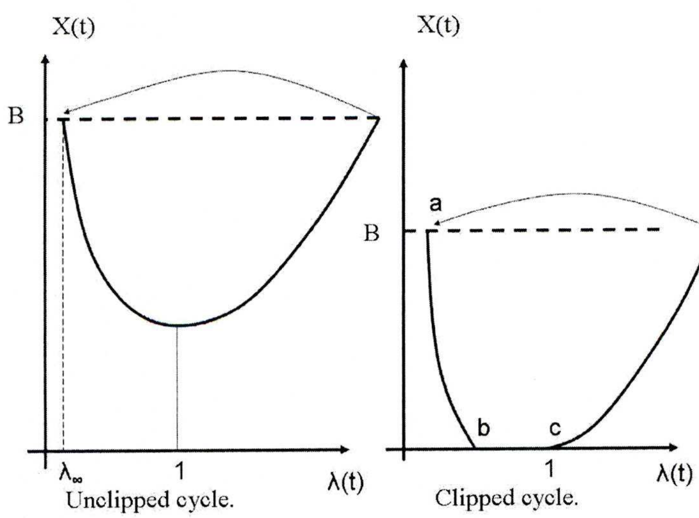

Figure 1: Possible cycles.

If $x(0) = B$, $\lambda(0) = \lambda_\infty$ is the starting point of a cycle then obviously $x(kT+0) = B$, $\lambda(kT+0) = \lambda_\infty$ for all integer $k \geq 0$.

It will be also shown that for fixed values of B and β, only one (simple) cycle exists which is stable in the following sense. Suppose $x(0)$ and $\lambda(0)$ are arbitrary and let λ_i be the value of $\lambda(t)$ immediately after the i-th jump. Then $\lim_{i\to\infty}\lambda_i = \lambda_\infty$. To put it different, any trajectory converges to the cycle which will be sometimes called 'limiting cycle'.

Main results

Theorem .1 *Let*

$$B_0 \triangleq \ln\frac{\beta-1}{\beta\ln\beta} - \frac{\beta\ln\beta}{1-\beta} - 1. \tag{3}$$

Then unclipped (clipped) cycle exists iff $B \geq B_0$ ($B < B_0$). The duration of the cycle equals $T = -\ln\beta$, for all $B \geq 0$.

B_0 denotes the minimal buffer size such that the queue is never empty, see Figure 1. Thus it is natural to call B_0 'critical' buffer size.

Remark .1 *If $r \neq 1$, $\mu \neq 1$ then the 'critical' buffer size will become $\tilde{B}_0 = (\mu/r)B_0$.*

Theorem .2 *Let λ_k be the value of $\lambda(t)$ immediately after the k-th jump. Then, starting from any initial value λ_0, $\lim_{k\to\infty}\lambda_k = \lambda_\infty$.*

(a) If $B \geq B_0$ then $\lambda_\infty = \frac{-\beta\ln\beta}{1-\beta}$.

(b) If $B < B_0$ then $\lambda_\infty = \beta e^\theta$, where θ is the single non-negative solution to $e^\theta - 1 - \theta = B$. In this case $\lambda_2 = \lambda_3 = \ldots = \lambda_\infty$.

(c) There exist only simple limiting cycles shown in figure 1, i.e. instant series of more than one jump are never realized, and all the values of $\lambda(t)$ immediately after a jump coincide with λ_∞ for the trajectory starting from $x_0 = B$, $\lambda_0 = \lambda_\infty$. The limiting cycle is stable.

Remark .2 *The meaning of variables can be understood from Figure 1: θ is the time interval corresponding to part c-d, δ is the time interval corresponding to the part a-b.*

Theorem .3 *If $B \geq B_0$ then*

$$\bar{\lambda} = 1; \quad \bar{x} = B + \frac{\ln\beta}{2} + \frac{\beta\ln\beta}{1-\beta} + 1.$$

If $B < B_0$ then

$$\bar{\lambda} = \frac{e^\theta(\beta-1)}{\ln\beta}; \quad \bar{x} = -\frac{\delta e^\theta(1-\beta) - \frac{1}{2}(\delta+\theta)^2}{\ln\beta},$$

where θ is defined in Theorem .2 and δ is the minimal positive solution to $\beta e^\theta e^\delta - \beta e^\theta - \delta + B = 0$.

Corollary .1 *Let $B < B_0$. Then θ, δ, $\bar{\lambda}$ and $\bar{x}$ monotonously increase with B. If B approaches zero then*

$$\bar{x} \to 0; \quad \bar{\lambda} \to \frac{\beta - 1}{\ln \beta}.$$

If B approaches B_0 then

$$\bar{x} \to \ln \frac{\beta - 1}{\beta \ln \beta} + \frac{\ln \beta}{2}; \quad \bar{\lambda} \to 1.$$

Now it is clear that the Pareto set for objectives $(\bar{x}, \bar{\lambda})$ is realized for $0 \leq B \leq B_0$: the minimal value of $\bar{x}$ equals zero and corresponds to $B = 0$; the maximal value of $\bar{\lambda}$ equals 1 and corresponds to $B = B_0$. If $B > B_0$ then $\bar{x}$ increases with B and $\bar{\lambda} = 1$ remains the same. Thus, solutions on the vertical dashed line (Figure 2), when $B > B_0$, are obviously dominated.

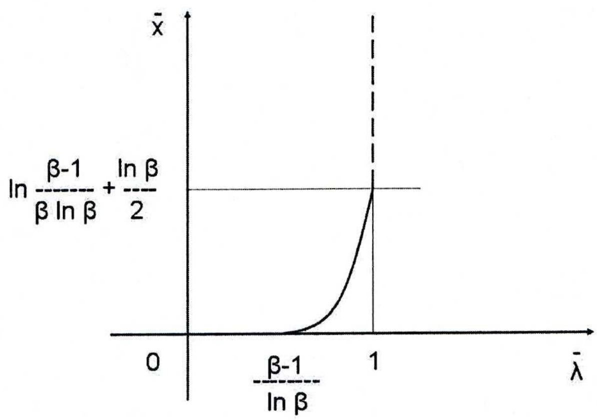

Figure 2: Tradeoff curve.

Proof of main statements

Proof of Theorem .1. Clearly, the first jump (or the first instant series of jumps) of the trajectory starting from $\lambda(0) = \lambda_0$, $x(0) = x_0$ results in the value $\lambda_1 \in [\beta, 1)$. Assuming the trajectory is not clipped and has no jumps on $[0, t]$, equations (1),(2) imply

$$\lambda(t) = \lambda_0 e^t, \quad x(t) = \lambda_0 e^t - t + x_0 - \lambda_0. \tag{4}$$

Equations $x_0 = x(T) = B$, $\beta\lambda(T) = \lambda_0$ result in formulae

$$T = \ln\frac{1}{\beta} \quad \text{(the duration of the cycle)}\,; \qquad \lambda_0 = \frac{-\beta\ln\beta}{1-\beta}.$$

The minimal value x_{min} of trajectory (4) starting from $(\lambda_0 = \frac{-\beta\ln\beta}{1-\beta}, x_0 = B)$ corresponds to $t^* = \ln\frac{1-\beta}{-\beta\ln\beta}$. Therefore an unclipped cycle exists iff

$$x_{min} = x(t^*) = B - B_0 \geq 0.$$

In what follows, we use denotation $\lambda_\infty = \frac{-\beta\ln\beta}{1-\beta}$ since the initial value λ_0 may be arbitrary.

In case $B < B_0$, starting from $(\lambda_0 = 1, x_0 = 0)$, the trajectory reaches level $x(\theta) = B$ at moment θ satisfying equation $e^\theta - 1 - \theta = B$, and initial values $(\lambda_0 = \beta e^\theta, x_0 = B)$ generate the trajectory with $x_{min} < 0$ meaning that we have constructed the clipped cycle. Conversely, if the clipped cycle exists then the above reasoning must lead to $x_{min} < 0$ which is equivalent to $B < B_0$. Note that the duration of the cycle equals $T = -\ln\beta$ for any $B \geq 0$. ■

Let us give some preliminary lemmas, based on which, the results in Theorem .2 become trivial.

Lemma .1 *If the constraint $x(t) \geq 0$ is withdrawn, for any $B \geq 0$, $\lim_{k\to\infty}\lambda_k = \lambda_\infty = \frac{-\beta\ln\beta}{1-\beta}$.*

Proof. For an arbitrary $\lambda_0 \in [\beta, 1)$ we define

$$\varphi(\lambda_0) = \beta\lambda_0 e^{T(\lambda_0)}, \tag{5}$$

where $T(\lambda_0)$ is the single positive solution to equation

$$F(\lambda_0, T) = \lambda_0 e^T - \lambda_0 - T = 0. \tag{6}$$

$T(\lambda_0)$ is well defined because of the observations of $F(\lambda_0, 0) = 0$, $F(\lambda_0, T) \to \infty$ as $T \to \infty$, $\left.\frac{\partial F}{\partial T}\right|_{T=0} = \lambda_0 - 1 < 0$, and $\frac{\partial^2 F}{\partial T^2} = \lambda_0 e^T > 0$.

Here $T(\lambda_0)$ is the time interval up to the next jump. This function $T(\lambda_0)$ is decreasing as well as $\varphi(\lambda_0)$. Indeed,

$$\frac{d\lambda_0(T)}{dT} = \frac{d\left(\frac{T}{e^T-1}\right)}{dT} = \frac{e^T - 1 - Te^T}{(e^T-1)^2} < 0,$$

$$\frac{dT(\lambda_0)}{d\lambda_0} = -\frac{\frac{\partial F}{\partial\lambda_0}}{\frac{\partial F}{\partial T}} = \frac{1-e^T}{\lambda_0 e^T - 1},$$

and

$$\begin{aligned}
\frac{d\varphi}{d\lambda_0} &= \frac{\partial\varphi}{\partial\lambda_0} + \frac{\partial\varphi}{\partial T}\frac{dT}{d\lambda_0} = \beta e^T + \beta\lambda_0 e^T \frac{1-e^T}{\lambda_0 e^T - 1} \\
&= \frac{\beta e^T(\lambda_0 e^T - 1) + \beta\lambda_0 e^T(1-e^T)}{\lambda_0 e^T - 1} \\
&= \frac{\beta e^T(\lambda_0 - 1)}{\lambda_0 e^T - 1} < 0.
\end{aligned}$$

One can show that $T(\beta) < -2\ln\beta$. Indeed, function $F(\beta,\cdot)$ has a single minimum and the value $F(\beta,-2\ln\beta)$ is already positive, because $\lim_{\beta\to 1} F(\beta,-2\ln\beta) = 0$ and $\frac{dF(\beta,-2\ln\beta)}{d\beta} = -\frac{(\beta-1)^2}{\beta^2} < 0$. Therefore $\varphi(\beta) = \beta^2 e^{T(\beta)} < 1$ meaning that in fact $\varphi : [\beta,\varphi(\beta)] \to [\beta,\varphi(\beta)]$. Another important consequence: starting from λ_1, instant series of more than one jump are never realized. (See Item (c).)

Since φ is decreasing, the double iteration $\psi(\lambda_0) \triangleq \varphi(\varphi(\lambda_0))$ is an increasing function meaning that the sequence $\lambda_2, \lambda_4, \lambda_6, \ldots$ is monotonous and hence converges to λ_∞ such that $\psi(\lambda_\infty) = \lambda_\infty$. We intend to prove that $\varphi(\lambda_\infty) = \lambda_\infty$. Suppose $\varphi(\lambda_\infty) = \lambda'_\infty \neq \lambda_\infty$ and let T_1 and T_2 be non-negative solutions to equations

$$\lambda_\infty e^{T_1} - \lambda_\infty - T_1 = 0, \quad \lambda'_\infty e^{T_2} - \lambda'_\infty - T_2 = 0. \tag{7}$$

We know that $\beta^2\lambda_\infty e^{T_1+T_2} = \lambda_\infty$, hence $\beta = e^{-(T_1+T_2)/2}$. Since

$$\lambda'_\infty = \varphi(\lambda_\infty) = \beta\lambda_\infty e^{T_1} = \beta e^{T_1}\frac{T_1}{e^{T_1}-1},$$

we have from (7): $\lambda'_\infty = \frac{T_2}{e^{T_2}-1} = \frac{\beta T_1 e^{T_1}}{e^{T_1}-1}$. Finally, using the formula for β we conclude that

$$\frac{T_1 e^{T_1/2}}{e^{T_1}-1} = \frac{T_2 e^{T_2/2}}{e^{T_2}-1}.$$

But function $j(\tau) = \frac{\tau e^{\tau/2}}{e^\tau - 1}$ is monotonous on $[0,\infty)$. Indeed, $\frac{dj}{d\tau} = \frac{e^{\frac{\tau}{2}}(e^\tau - \frac{\tau}{2}e^\tau - 1 - \frac{\tau}{2})}{(e^\tau-1)^2}$, where the numerator takes the value of zero when $\tau = 0$, which, in addition to $\frac{d}{d\tau}(e^\tau - \frac{\tau}{2}e^\tau - 1 - \frac{\tau}{2}) = \frac{1}{2}(e^\tau - \tau e^\tau - 1) < 0$, implies that $g(\tau)$ is strictly decreasing. Hence $T_1 = T_2$ and $\lambda_\infty = \lambda'_\infty$.

Equation $\varphi(\lambda_\infty) = \lambda_\infty$ results in formulae

$$T = \ln\frac{1}{\beta} \quad \text{(the duration of the cycle)}\,; \qquad \lambda_\infty = \frac{-\beta\ln\beta}{1-\beta},$$

and no other cycle exists. ■

Remark .3 *Consider any fixed $\lambda_0 \in [\beta, 1)$ and $\varphi(\cdot)$ defined in Lemma .1. When the constraint $x(t) \geq 0$ is withdrawn, $\lambda_1 = \varphi(\lambda_0)$. With the constraint $x(t) \geq 0$, $\varphi(\lambda_0) = \lambda_1$*

if the trajectory starting from $(\lambda_0, x_0 = B)$ is unclipped.

Lemma .2 *If the constraint $x(t) \geq 0$ is withdrawn, for any arbitrary $\lambda_n \in [\beta, 1)$, $\lambda_{n+2} \in [\min\{\lambda_n, \lambda_{n+1}\}, \max\{\lambda_n, \lambda_{n+1}\}]$.*

Proof. For any $\lambda_0 \in [\beta, 1)$, there are three possibilities about λ_1: $\lambda_1 < \lambda_0$, $\lambda_1 = \lambda_0$, or $\lambda_1 > \lambda_0$.

Consider the case $\lambda_1 < \lambda_0$. Then there are two possibilities about λ_2: $\lambda_2 \in [\lambda_1, \lambda_0]$ or $\lambda_2 > \lambda_0 > \lambda_1$. Clearly one only needs to consider the possibility of $\lambda_2 > \lambda_0 > \lambda_1$, which, together with Remark .3 and the fact $\varphi()$ is decreasing and $\psi(\cdot)$ is increasing, where $\varphi(\cdot)$ and $\psi(\cdot)$ are defined in Lemma .1, leads to $\forall n \in \{3, 4 \ldots\}$ $\lambda_{2n} > \lambda_{2(n-1)} > \cdots > \lambda_2 > \lambda_0 > \lambda_1 > \cdots > \lambda_{2n-1} > \lambda_{2n+1}$. But this contradicts Lemma .1.

The case of $\lambda_0 < \lambda_1$ can be treated in exactly the same way, and the case of $\lambda_0 = \lambda_1$ is trivial. ■

Lemma .3 *If the trajectory starting with $\lambda_0 \in [\beta, 1)$ is clipped, then there exists a single $\hat{\lambda} \in (\lambda_0, 1)$, starting with which the trajectory is critical, by which is meant that the trajectory only has its x component being zero at one time point.*

Proof. Suppose the constraint $x(t) \geq 0$ is withdrawn. Then the trajectory starting with λ_0 attains its minimum at $\hat{t}_{\lambda_0} = \ln \frac{1}{\lambda_0}$ such that $\lambda(\hat{t}_{\lambda_0}) = 1$, and $x(\hat{t}_{\lambda_0}) = 1 - \ln \frac{1}{\lambda_0} + B - \lambda_0 \triangleq \hat{x}(\lambda_0, B)$. (See (4).) Clearly, $\frac{\partial \hat{x}}{\partial \lambda_0} > 0$.

Now impose the constraint $x(t) \geq 0$. That the trajectory starting with λ_0 is clipped (critical) is equivalent to $\hat{x}(\lambda_0, B) < 0$ $(= 0)$. Then the lemma follows from the fact that $\hat{x}(1, B) = B > 0$, $\hat{x}(\lambda_0, B)$ is increasing and continuous in λ_0. ■

Remark .4 *It can be the case that $\hat{\lambda}$ fails to exist, for example when B is sufficiently large. However, it follows from the proof of Lemma .3 that when $\hat{\lambda}$ exists, trajectories starting with $\lambda_0 \in (\hat{\lambda}, 1)$ are unclipped.*

Remark .5 *With the constraint $x(t) \geq 0$, Lemma .2, Lemma .3 and Remark .4 imply that once two consecutive unclipped trajectories are realized, all the further trajectories will be unclipped. In this case, the limiting cycle, whose existence follows now from Lemma .1, will be unclipped.*

Proof of Theorem .2 (a) Consider case $B \geq B_0$. Fix some arbitrary $\lambda_0 \in [\beta, 1)$. If the trajectory starting with λ_0 is clipped, then $\hat{\lambda}$ is well defined, and the one starting from λ_1 cannot be clipped, since otherwise it would contradict Theorem .1. Thus we have $\lambda_1 > \hat{\lambda}$. Furthermore $\lambda_2 \in [\min\{\lambda_1, \hat{\lambda}\}, \max\{\lambda_1, \hat{\lambda}\}]$ due to Lemma .2. We now already have two consecutive unclipped trajectories and the result follows from Remark .5. If the trajectories starting with λ_0 and λ_1 are unclipped, then the result follows from Remark .5 and Lemma .1. If the trajectory starting with λ_0 is unclipped, but the one starting with λ_1 is clipped, we have $\hat{\lambda} \in [\beta, 1)$ well defined. In addition, $\lambda_2 = \varphi(\hat{\lambda}) \geq \hat{\lambda}$ since otherwise a clipped cycle will be realized, contradicting Theorem .1. Then $\lambda_3 = \varphi(\lambda_2) \in [\min\{\lambda_2, \hat{\lambda}\}, \max\{\lambda_2, \hat{\lambda}\}]$ due to Lemma .2. Hence now we have two consecutive unclipped trajectories and the result follows from Remark .5.

Since

$$\frac{d\varphi}{d\lambda_0} = \frac{\beta e^T (T - e^T + 1)}{Te^T - e^T + 1},$$

where T is given by (6), we conclude that at the stable point

$$\left.\frac{d\varphi}{d\lambda_0}\right|_{\lambda_\infty} = \frac{\beta - 1 - \beta \ln \beta}{\beta - 1 - \ln \beta} > -1$$

meaning that $\left|\frac{d\varphi}{d\lambda_0}\right| < 1$ in the neighborhood of λ_∞, so that the limiting cycle is stable.

(b) If $B < B_0$ then, maximum after one jump (starting from λ_1), the trajectory is clipped, because otherwise we would have faced a sequence $\lambda_1, \lambda_2, \ldots$ resulting in unclipped trajectories and, according to Remark .5, there would have existed an unclipped limiting cycle. Therefore, $\lambda_2 = \lambda_3 = \ldots = \lambda_\infty$, and the value $\lambda_\infty = \beta e^\theta$ follows from equation $x(\theta) = e^\theta - 1 - \theta = B$ describing the trajectory starting from $(\lambda_0 = 1, x_0 = 0)$.

(c) The last statement follows from the previous reasoning. ∎

Proof of Theorem .3. According to Theorems .1, .2, the trajectories converge to a single cycle with initial values $(\lambda_0 = \lambda_\infty, x_0 = B)$. Hence

$$\bar{\lambda} = \frac{1}{T}\int_0^T \lambda(t)dt, \quad \bar{x} = \frac{1}{T}\int_0^T x(t)dt.$$

Consider the case $B \geq B_0$. Having in mind Theorem .2, formulae (4) and $T = -\ln\beta$. We have

$$\bar{\lambda} = \frac{1}{T}\int_0^T \lambda_\infty e^s ds = \frac{1}{T}\lambda_\infty e^s\Big|_0^T = \frac{1}{\ln\frac{1}{\beta}}\left(\frac{\ln\frac{1}{\beta}}{1-\beta} - \frac{\beta\ln\frac{1}{\beta}}{1-\beta}\right) = 1,$$

and

$$\begin{aligned}
\bar{x} &= \frac{1}{T}\int_0^T x(s)ds = \frac{1}{T}\int_0^T [\lambda(s) - s + B - \lambda_\infty]ds \\
&= \frac{1}{T}\int_0^T \lambda(s)ds - \frac{T}{2} + B - \frac{\beta \ln\frac{1}{\beta}}{1-\beta} \\
&= B + \frac{\ln\beta}{2} + \frac{\beta\ln\beta}{1-\beta} + 1.
\end{aligned}$$

Consider the case $B < B_0$. Here one has to calculate the integrals along branches a-b, b-c and c-d (Figure 1). Additionally let us put $\eta \triangleq T - \theta$. Equation for δ comes from condition $x(\delta) = 0$ for trajectory starting from $(\lambda_0 = \lambda_\infty = \beta e^\theta, x_0 = B)$. Before further calculations for $\bar{\lambda}$ and $\bar{x}$, let us verify that parameters θ, η and δ are well defined.

For θ, according to (4), θ satisfies the equation $B = e^{\theta} - 1 - \theta$, which has a single positive solution since on $[0,\infty)$ the right hand side increases with θ from zero up to ∞.

For δ, again according to (4), δ satisfies the equation

$$\begin{aligned} &x(\delta) = B + \lambda(\delta) - \lambda(0) - \delta = B + \beta e^{\theta} e^{\delta} - \beta e^{\theta} - \delta = 0 \\ \Rightarrow \quad & \beta e^{\theta} e^{\delta} - \delta = \beta e^{\theta} - B. \end{aligned} \tag{8}$$

In the below we will verify that the above equation has two positive solutions.

Let us put $g(\delta) = \beta e^{\theta} e^{\delta} - \delta$, with $\frac{dg}{d\delta} = \beta e^{\theta} e^{\delta} - 1$, and $\left.\frac{dg}{d\delta}\right|_{\delta=0} = \beta e^{\theta} - 1$. Since $g(0) = \beta e^{\theta} > \beta e^{\theta} - B$, in order for solutions to (8) to exist, $\frac{dg}{d\delta}$ must be negative at $\delta = 0$. Otherwise, $g(\delta)$ will be increasing, never meeting $\beta e^{\theta} - B$. Furthermore, we must also have namely $g_{min} \leq \beta e^{\theta} - B$, where g_{min} means the minimum of $g(\delta)$ on $[0,\infty)$, which clearly exists. In particular, in case $g_{min} = \beta e^{\theta} - B$, there will be two identical solutions to (8).

Firstly, let us verify that $g_{min} - \beta e^{\theta} + B \leq 0$. One can check that $g_{min} = 1 + \ln\beta + \theta$. Keeping in mind that the β-independent θ solves $e^{\theta} = 1 + \theta + B$ and increases with B, we have $g_{min} - \beta e^{\theta} + B = e^{\theta}(1-\beta) + \ln\beta$. Also notice that $\theta \leq \theta_{max}$, where θ_{max} is the single positive solution to $1 + \theta_{max} + B_0 = e^{\theta_{max}}$. It then suffices to show $e^{\theta_{max}}(1 - \beta) + \ln\beta \triangleq g_1(\theta_{max}) \leq 0$. Recall (3) for B_0. Then θ_{max} solves

$$\begin{aligned} &\theta_{max} + \ln\frac{1-\beta}{\beta\ln\frac{1}{\beta}} + \frac{\beta\ln\frac{1}{\beta}}{1-\beta} = e^{\theta_{max}} \\ \Leftrightarrow \quad & \theta_{max} - e^{\theta_{max}} - \ln\left(\frac{\ln\frac{1}{\beta}}{1-\beta}\right) + \frac{\ln\frac{1}{\beta}}{1-\beta} = 0. \end{aligned}$$

The observation of $e^{\ln\left(\frac{\ln\frac{1}{\beta}}{1-\beta}\right)} = \frac{\ln\frac{1}{\beta}}{1-\beta}$ indicates that $\theta_{max} = \ln\left(\frac{\ln\frac{1}{\beta}}{1-\beta}\right)$. Now we have $g_1(\theta_{max}) = 0$ as required.

Secondly, let us verify $\left.\frac{dg}{d\delta}\right|_{\delta=0} = \beta e^{\theta} - 1 < 0$. Equivalently, we will show $e^{\theta} < \frac{1}{\beta}$. From the above proof we already know that $e^{\theta} \leq e^{\theta_{max}} = \frac{\ln\frac{1}{\beta}}{1-\beta}$. However, it can be easily checked that $\frac{\ln\frac{1}{\beta}}{1-\beta} < \frac{1}{\beta}$, as required. Hence, we conclude that $\beta e^{\theta} e^{\delta} - \beta e^{\theta} - \delta + B = 0$ has no more than two positive solutions, and δ is corresponding to the smaller one.

For η to be well defined, we need to show $T - \theta = \frac{1}{\beta} - \theta \geq 0$. But we already know in the above proof that $e^{\theta} \leq e^{\theta_{max}} < \frac{1}{\beta} = e^{T}$ as required. Up to now, we have verified that all the parameters θ, δ and η are well defined, and we are ready to calculate $\bar{\lambda}$ and $\bar{x}$.

For $\bar{\lambda}$, with in mind $T = \ln\frac{1}{\beta}$ and (4), we have $\bar{\lambda} = \frac{1}{T}\int_0^T \lambda(s)ds = \frac{1}{T}\lambda_{\infty}(e^{T} - 1) =$

$\frac{\beta e^{\theta}(\frac{1}{\beta}-1)}{\ln\frac{1}{\beta}} = \frac{e^{\theta}(\beta-1)}{\ln\beta}$.

For $\bar{x}$, to get the desired expression, we will do some manipulations firstly. Due to (4), we have

$$\lambda_\infty = \frac{1}{e^\eta} = \beta e^\theta, \tag{9}$$

and

$$\begin{aligned} & x(\delta) = B + \lambda(\delta) - \lambda_\infty - \delta = 0 \\ \Leftrightarrow \quad & \lambda_\infty = \frac{\delta - B}{e^\delta - 1} = \frac{\delta - e^\theta + 1 + \theta}{e^\delta - 1}, \end{aligned} \tag{10}$$

where we have replaced B by $e^\theta - 1 - \theta$. (See Theorem .2 for the definition of θ.)

Now equate (9) to (10), and we have

$$e^\delta - 1 = e^\eta(\delta - e^\theta + 1 + \theta). \tag{11}$$

On one hand,

$$\begin{aligned} \int_0^\delta x(s)ds &= B\delta + \lambda_\infty e^\delta - \lambda_\infty - \lambda_\infty\delta - \frac{\delta^2}{2} \\ &= e^\theta\delta - \delta - \theta\delta + \frac{e^\delta - 1}{e^\eta} - \frac{\delta}{e^\eta} - \frac{\delta^2}{2} \\ & \quad \text{(by using the definition of } \theta \text{ and (9))} \\ &= e^\theta\delta - \theta\delta - e^\theta + 1 + \theta - \frac{\delta}{e^\eta} - \frac{\delta^2}{2} \text{ (by using (11)).} \end{aligned}$$

On the other hand,

$$\begin{aligned} \int_\eta^{\eta+\theta} x(s)ds &= \int_0^\theta (x(\eta) + e^s - 1 - s)ds \text{ (due to (4))} \\ &= e^\theta - 1 - \theta - \frac{\theta^2}{2}. \end{aligned}$$

Eventually, we have

$$\begin{aligned} \bar{x} &= \frac{1}{T}\left(\int_0^\delta x(s)ds + \int_\eta^{\eta+\theta} x(s)ds\right) \\ &= \frac{1}{T}\left(e^\theta\delta - \theta\delta - \frac{\delta}{e^\eta} - \frac{\delta^2}{2} - \frac{\theta^2}{2}\right) \\ &= -\frac{\delta e^\theta(1-\beta) - \frac{1}{2}(\delta+\theta)^2}{\ln\beta} \text{ (using } T = \ln\tfrac{1}{\beta} \text{ and (4)).} \end{aligned}$$

■

Proof of Corollary .1. The monotonicity of θ follows directly from its equation. The value of $\bar{\lambda}$ increases because θ increases. δ is the first moment when coordinate $x(t)$ becomes zero. As B and $\lambda_\infty = \beta e^\theta$ increase, the value $x(t)$ at an arbitrary t corresponding to the part a-b (Figure 1) increases with B meaning that δ increases. The reasoning presented implies that the value of the integral $\int_0^T x(t)dt$ increases, and hence $\bar{x}$ increases. (Remember $T = -\ln\beta$ is constant.) To be more precise, by combining (9) and (10) and using the definition of θ we have

$$\beta e^\theta e^\delta - \beta e^\theta - \delta + e^\theta - 1 - \theta = 0 \tag{12}$$
$$\Rightarrow \quad \frac{d\delta}{d\theta} = -\frac{\beta e^\theta e^\delta - \beta e^\theta + e^\theta - 1}{\beta e^\theta e^\delta - 1},$$

where the numerator is obviously positive, and the denominator is negative, meaning $\frac{d\delta}{d\theta} > 0$. Indeed, $\beta e^{\theta+\delta} < 1$ follows from the fact $\theta + \delta < T = \ln\frac{1}{\beta}$, when $B < B_0$. Recall that $\frac{d\theta}{dB} > 0$, we eventually have $\frac{d\delta}{dB} > 0$.

For $\frac{d\bar{x}}{dB}$, we firstly have

$$\begin{aligned}
\frac{d\ln\frac{1}{\beta}\bar{x}}{d\theta} &= \delta e^\theta(1-\beta) + (1-\beta)e^\theta\frac{d\delta}{d\theta} - (\theta+\delta)(\frac{d\delta}{d\theta} + 1) \\
&= \frac{d\delta}{d\theta}(e^\theta - \beta e^\theta - \delta - \theta) + \delta e^\theta(1-\beta) - \delta - \theta \\
&= \beta e^\theta e^\delta - \beta e^\theta + e^\theta - 1 + \delta e^\theta - \delta e^\theta\beta - \delta - \theta \text{ (due to (12))} \\
&= \delta e^\theta(1-\beta) > 0 \text{ (due to (12) again)}.
\end{aligned}$$

Hence $\frac{d\bar{x}}{dB} > 0$. (Recall $\frac{d\theta}{dB} > 0$.)

If $B \to 0$ then $\theta \to 0$ and $\delta \to 0$. If $B \to B_0$ then

$$\theta \to \ln\left(\frac{-\ln\beta}{1-\beta}\right); \quad \delta \to T - \theta = -\ln\beta - \ln\left(\frac{-\ln\beta}{1-\beta}\right) = -\ln\left(\frac{-\beta\ln\beta}{1-\beta}\right).$$

Calculations of the limits are straightforward. ■

Bibliography

Aliprantis C, Border K (2007) Infinite Dimensional Analysis-a Hitchhiker's Guide, 3rd edn. Springer, NY

Allman M, Paxson V, Stevens V (1999) TCP Congestion Control. RFC 2581

Altman E (1999) Constrained Markov Decision Processes. Chapman and Hall/CRC, Boca Raton

Altman E, Jimenez T, Koole G (2001) On the Comparison of Queueing Systems with their Fluid Limits. Probab Engrg and Inform Sci 15:165–178

Altman E, Avrachenkov K, Prabhu B (2005) Fairness in MIMD Congestion Control Algorithms. In: Proc. of IEEE INFOCOM

Appenzeller G, Keslassy I, McKeown N (2004) Sizing Router Buffers. In: Proc. of ACM SIGCOMM

Avrachenkov K, Ayesta U, Altman E, Nain P, Barakat C (2002) The Effect of Router Buffer Size on The TCP Performance. In: Proc. of LONIIS workshop, Russia

Avrachenkov K, Ayesta U, Piunovskiy A (2005) Optimal Choice of Buffer Size of an Internet Router. In: Proc. of IEEE CDC/ECC, Spain

Avrachenkov K, Ayesta U, Piunovskiy A (2010) Convergence and Optimal Buffer Sizing for AIMD Congestion Control. Perform Evaluation Doi:10.1016/j.peva.2010.01.004

Ayesta U, Piunovskiy A, Zhang Y (2008) Fluid Model of An Internet Router Under the MIMD Control Scheme. In: Raghavan S, Golden B, Wasil E (eds) Telecommunications Modeling, Policy and Technology, Springer-Verlag, New York, pp 239–251

Baker R, Urban T (1988) A Determinisitc Inventory System with an Inventory-Level-Dependent Demand Rate. J Oper Res Soc 39(9):823–831

Bäuerle N (2000) Asymptotic Optimality of Tracking Policies in Stochastic Networks. Ann Appl Prob 10(4):1065–1083

Bäuerle N (2002) Optimal Control of Queueing Networks: an Approach via Fluid Models. Adv Appl Prob 34:313–328

Berman O, Perry D (2006) An EOQ Model with State-Dependent Demand Rate. Eur J Oper Res 171:255–272

Berman O, Perry D, Stadje W (2007) Performance Analysis of a Fluid Production/Inventory Model with State-Dependence. Methodol Comput Appl Probab 9:465–481

Bertsekas D, Shreve S (1978) Stochastic Optimal Control: the Discrete-Time Case. Academic Press, NY

Bohacek S, Hespanha J, Lee J, Obraczka K (2003) A Hybrid Systems Modeling Framework for Fast and Accurate Simulation of Data Communication Networks. In: Proc. of ACM SIGMETRICS

Bolot J, Shankar A (1992) Analysis of a Fluid Approximation to Flow Control Dynamics. In: Proc. of IEEE INFOCOM

Bremaud P (1999) Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues. Springer, NY

Cantrell P (1986) Computation of the Transient M/M/1 Queue CDF, PDF, and Mean with Generalized Q-Functions. IEEE Transactions on Communications COM-34(8)

Chen H (1995) Fluid Approximations and Stability of Multiclass Queueing Networks: Work-Conserving Disciplines. Ann Appl Prob 5(3):637–665

Chen H (1996) Rate of Convergence of the Fluid Approximation for Generalized Jackson Networks. J Appl Probab 33:804–814

Chen H, Mandelbaum A (1991) Discrete Flow Networks: Bottleneck Analysis and Fluid Approximations. Math Oper Res 16(2):408–446

Chen H, Mandelbaum A (1994) Hierachical Modeling of Stochastic Networks, Part I: Fluid Models. In: Yao D (ed) Stochastic Modeling and Analysis of Manufacturing Systems, Springer-Verlag, New York, pp 47–105

Clancy D, Piunovskiy A (2005) An Explicit Optimal Isolation Policy for a Determinisitc Epidemic Model. Applied Mathematics and Computation 163:1109–1121

Dai J (1995) On Positive Harris Recurrence of Multiclass Queueing Networks: A Unified Approach via Fluid Limit Models. Ann Appl Prob 5:49–77

Datta T, Pal A (1990) A Note on an Inventory Model with Inventory-Level-Dependent Demand Rate. J Oper Res Soc 41(10):971–975

Dhamdhere A, Dovrolis C (2006) Open issues in Router Buffer Sizing. ACM SIGCOMM Computer Communication Review 36(1):87–92

Dhamdhere A, Jiang H, Dovrolis C (2005) Buffer Sizing for Congested Internet Links. In: Proc. of IEEE INFOCOM

Dorini G, Di Pierro F, Savic D, Piunovskiy A (2007) Neighbourhood Search for Constructing Pareto Sets. Math Meth Oper Res 65

Dynkin E, Yushkevich A (1979) Controlled Markov Processes. Springer-Verlag, NY

Enachescu M, Ganjali Y, Goel A, McKeown N (2005) Part III: Routers with Very Small Buffers. ACM SIGCOMM CCR 35(2):83–89

Ethier S, Kurtz T (1986) Markov Processes: Chracterization and Convergence. Wiley, NY

Fall K, Floyd S (1996) Simulation-based Comparisons of Tahoe, Reno and SACK TCP. Computer Communications Review 26:5–21

Feinberg E (2004) Continuous Time Discounted Jump Markov Decision Processes: a Discrete-Event Approach. Math Oper Res 29(3):492–524

Ferguson M, Jayaraman V, Souza G (2007) Note: An Application of the EOQ Model with Nonlinear Holding Cost to Inventory Management of Perishables. Eur J Oper Res 180:485–490

Floyd S (2003) HighSpeed TCP for Large Congestion Windows. RFC 3649

Foss S, Kovalevskii A (1999) A Stability Criterion via Fluid Limits and its Application to a Polling System. Queueing Syst 32:131–168

Gajrat AS, Hordijk A (2000) Fluid Approximations of a Controlled Multiclass Tandem Network. Queueing Syst 35:349–380

Gajrat AS, Hordijk A (2005) On the Structure of the Optimal Server Control for Fluid Networks. Math Meth Oper Res 62:55–75

Gajrat AS, Hordijk A, Malyshev V, Spieksma F (1997) Fluid Approximations of Markov Decision Chains. Markov Process Related Fields 3:129–150

Gajrat AS, Hordijk A, Ridder A (2003) Large-Deviations Analysis of the Fluid Approximations for a Controllable Tandem Queue. Ann Appl Prob 13(4):1423–1448

Ganjali Y, McKeown N (2006) Update on Buffer Sizing in Internet Routers. ACM Computer Communication Review 36(5):67–70

Gemignani M (1990) Elementary Topology. Dover, NY

Gikhman I, Skorokhod A (1996) Introduction to the Theory of Random Processes. Dover, NY

Giri B, Chaudhuri K (1998) Deterministic Models of Perishable Inventory with Stock-Dependent Demand Rate and Nonlinear Holding Cost. Eur J Oper Res 105:467–474

Giri B, Pal S, Goswami A, Chaudhuri K (1996) An Inventory Model for Deteriorating Items with Stock-Dependent Demand Rate. Eur J Oper Res 95:604–610

Gleissner W (1988) The Spread of Epidemics. Appl Math Comput 27:167–171

Gorinsky S, Kantawala A, Turner J (2005) Link Buffer Sizing: A New Look at the Old Problem. In: Proc. of ISCC

Guo X (2007) Continuous-Time Markov Decision Processes with Discounted Rewards: the Case of Polish Spaces. Math Oper Res 32(1):73–87

Guo X, Dai Y (2002) The Unbounded Cost Discounted Model for Continuous Time Markov Decision. Acta Mathematica Sinica 45(1):171–182, in Chinese

Guo X, Hernández-Lerma O (2003a) Constrained Continuous-Time Markov Control Processes with Discounted Criteria. Stoch Ana Appl 21(2):379–399

Guo X, Hernández-Lerma O (2003b) Continuous-Time Controlled Markov Chains. Ann Appl Prob 13:363–388

Guo X, Hernández-Lerma O (2003c) Continuous-Time Controlled Markov Chains with Discounted Rewards. Acta Appl Math 79:195–216

Guo X, Hernández-Lerma O (2003d) Drift and Monotonicity Condtions for Continuous-Time Controlled Markov Chains with an Average Criterion. IEEE Trans Automat Control 48(2):236–245

Guo X, Hernández-Lerma O (2009) Continuous-Time Markov Decision Processes: Theory and Applications. Springer, Heidelberg

Guo X, Piunovskiy A (2010) Discounted Continuous-Time Markov Decision Processes with Constraints: Unbounded Transition and Loss Rates, submitted to Math Oper Res

Guo X, Rieder U (2006) Average Optimality for Continuous-Time Markov Decision Processes in Polish Spaces. Ann Appl Prob 16(2):730–756

Guo X, Zhu W (2002) Denumerable-State Continuous-Time Markov Decision Processes with Unbounded Transition and Reward Rates under the Discounted Criterion. J Appl Probab 39(2):233–250

Guo X, Hernández-Lerma O, Prieto-Rumeau T (2006) A Survey of Recent Results on Continuous-Time Markov Decision Processes. Top 14:177–257

He Q, Jewkes E (2002) Performance Measures of a Make-to-Order Inventory-Production System. IIE Trans 32:409–419

Hernández-Lerma O (1994) Lectures on Continuous-Time Markov Control Processes. Sociedad Matematica Mexicana, Mexico city

Hernández-Lerma O, Lasserre J (1996) Discrete-Time Markov Control Processes-Basic Optimality Criterion. Springer-Verlag, NY

Hernández-Lerma O, Lasserre J (1999) Further Topics on Discrete-Time Markov Control Processes. Springer-Verlag, NY

Hespanha J, Bohacek S, Obraczka K, Lee J (2001) Hybrid Modeling of TCP Congestion Control. Hybrid Systems: Commuputation and Control, LNCS 2034:291–304

Hu Q, Liu J (1999) An Introduction to Markov Decision Processes. Xidian University Press, Xi'an, in Chinese

Jacobson S (1988) Congestion Avoidance and Control. In: Proc. of SIGCOMM

Jacod J (1975) Multivariate Point Processes: Predictable Projection, Radon-Nykodym Derivatives, Representation of Martingales. Z Wahrscheinlichkeitstheorie verw Gebiete 31:235–253

Jin C, Wei D, Low S (2004) FAST TCP: Motivation, Architecture, Algorithms, Performance. In: Proc. of IEEE INFOCOM

Kelly T (2003) Scalable TCP: Improving Performance in High-Speed Wide Area Networks. Computer Communication Review 33(2):83–91

Kermack W, McKendrick A (1927) A Contribution to the Mathematical Theory of Epidemics. Proc Roy Soc Lond A 115:700–721

Khoury R, Altman E (2004) Analysis of scalable TCP. In: Proc. of HET-NET, UK

Kitaev M (1986) Semi-Markov and Jump Markov Controlled Models: Average Cost Criterion. Theory Probab Appl 30(2):272–288

Kitaev M, Rykov V (1995) Controlled Queueing Systems. CRC Press, Boca Raton

Köchel P (1996) On Queueing Models for some Multi-Location Problems. Int J Prod Econ 45:429–433

Lakshmikantha A, Srikant R, Beck C (2008) Impact of File Arrivals and Departures on Buffer Sizing in Core Routers. In: Proc. of IEEE INFOCOM

Leith D, Shorten R (2004) H-TCP: TCP for High Speed and Long-Distance Networks. In: Proc. of PFLDnet

Levin R, McLaughlin C, Lamone R, Kottas J (1972) Productions/Operations Management: Contemporary Policy for Managing Operating Systmes. McGraw-Hill, New York

Li D, Glazebrook K (2010) An Approximate Dynamic Programming Approach to the Development of Heuristics for the Scheduling of Impatient Jobs in a Clearing System. Nav Res Log Doi 10.1002/nav.20395

Li Y, Leith D, Shorten R (2007) Experimental Evaluation of TCP Protocols for High-Speed Networks. IEEE ACM T Network 15(5):1109–1122

Maglaras C (2000) Discrete-Review Policies for Scheduling Stochastic Networks: Trajectory Tracking and Fluid-Scale Asymptotic Optimality. Ann Appl Prob 10(3):897–929

Mandelbaum A, Pats G (1995) State-dependent Queues: Approximations and Applications. In: Kelly F, Williams R (eds) Stochastic Networks, Springer, NY, pp 239–282, proc. of the IMA

Mandelbaum A, Massey W, Reiman M (1998) Strong Approximations for Markovian Service Networks. Queueing Syst 30:149–201

Massoulié L, Roberts J (2000) Bandwidth Sharing and Admission Control for Elastic Traffic. Telecommun Syst 15:185–201

Meyn S (1997) Stability and optimization of Queueing Networks and Their Fluid Models. In: Yin G, Zhang Q (eds) Mathematics of Stochastic Manufacturing Systems (Lectures Appl Math 33), American Mathematical Society, Providence, pp 175–199

Meyn S (2005) Dynamic Safety-Stocks for Asymptotic Optimality in Stochastic Networks. Queueing Syst 50(2-3):255–297

Morris R (1997) TCP Behaviour with Many Flows. In: Proc. of IEEE International Conference on Network Protocols

Morris R (2000) Scalable TCP Congestion Control. In: Proc. of IEEE INFOCOM

Pang G, Day M (2007) Fluid Limits of Optimally Controlled Queueing Networks. J Appl Math Stoch Anal Article ID 68958, doi:10.I155/2007/68958

Piunovskiy A (1997) Optimal Control of Random Sequences in Problems with Constraints. Kluwer Academic Publishers, Dordrecht

Piunovskiy A (1998) A Controlled Jump Discounted Model with Constraints. Theory Probab Appl 42(1):51–71

Piunovskiy A (2004a) Bicriteria Optimization of a Queue with a Controlled Input Stream. Queueing Syst 48(1-2):159–184

Piunovskiy A (2004b) Multicriteria Impulsive Control of Jump Markov Processes. Math Meth Oper Res 60:125–144

Piunovskiy A (2005) Discounted Continuous Time Markov Decision Processes: the Convex Analytic Approach. In: Proc. of 16th IFAC World Congress, Prague

Piunovskiy A (2009a) Controlled Jump Markov Processes with Local Transitions and Their Fluid Approximation. WSEAS Trans on Systems and Control 4(8):399–412

Piunovskiy A (2009b) Random Walk, Birth-and-Death Process and Their Fluid Approximations: 'Absorbing Case'. Math Meth Oper Res 70(2):285–312

Piunovskiy A, Clancy D (2008) An Explicit Optimal Intervention Policy for a Deterministc Epidemic Model. Optim Control Appl Meth 29:413–428

Prasad R, Dovrolis C, Thottan M (2007) Router Buffer Sizing Revisited: The Role of The Output/Input Capacity Ratio. In: Proc. of ACM SIGCOMM CoNEXT

Pullan M (1995) Forms of Optimal Solutions for Separated Continuous Linear Programs. SIAM J Control Optimization 33:1952–1977

Puterman M (1994) Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley Series in Probability and Mathematical Statistics, Wiley, NY

Raina G, Wischik D (2005) Buffer Sizes for Large Multiplexers: TCP Queueing Theory and Instability Analysis. In: Proc. of EuroNGI

Raina G, Towsley D, Wischik D (2005) Part II: Control Theory for Buffer Sizing. ACM SIGCOMM CCR 35(2):79–82

Revuz D, Yor M (1999) Continuous Martingales and Brownian Motion, 3rd edn. Grundlehren der mathematischen Wissenschaften, Vol. 293, Springer

Rhee I, Xu L (2005) CUBIC: A New TCP-friendly High-Speed TCP Variant. In: Proc. of PFLDnet

Robert P (2003) Stochastic Networks and Queues, English edn. 52 Applications of Mathematics, Springer, NY

Rockafellar R (1974) Conjugate Duality and Optimization. CBMS 16, SIAM, Philadelphia

Ross S (2002) Introduction to Probability Models, 8th edn. Academic press

Rothblum U, Shorten R (2007) Nonlinear AIMD Congestion Control and Contraction Mappings. SIAM J Control Optim 46(5):1882–1896

Schäl M (1975) On Dynamic Programming: Compactness of the Space of Policies. Stoch Proc Appl 3:345–364

Sharma O, Tarabia A (2000) A Simple Transient Analysis of an M/M/1/N Queue. The Indian Journal of Statistics 62(2):273–281

Stanojević R, Shorten R, Kellett C (2006) Adaptive Tuning of Drop-Tail Buffers for Reducing Queueing Delays. IEEE Commun Lett 10(7):570–572

Tan B, Ying L, Srikant R (2010) Short-Term Fairness and Long-Term QoS in the Internet. Perform Evaluation 67:406–414

Urban T (1992) An Inventory Model with an Inventory-Level-Dependent Demand Rate and Relaxed Terminal Conditions. J Oper Res Soc 43(7):721–724

Urban T (1995) Inventory Models with the Demand Rate Dependent on Stock and Shortage Levels. Int J Prod Econ 40:21–28

Urban T (2005) Inventory Models with Inventory-Level-Dependent Demand: A Comprehensive Review and Unifying Theory. Eur J Oper Res 162:792–804

Verloop I (2009) Scheduling in Stochastic Resource-Sharing Systems. PhD thesis, Eindhoven University of Technology

Verloop I, Núñez Queija R (2007) Efficient Resource Allocation in Bandwidth-Sharing Networks. ACM SIGMETRICS Performance Evaluation Review 35(3):49–50

Villamizar C, Song C (1994) High Performance TCP in ANSNET. ACM Computer Communication Review 24(5):44–60

Vu-Brugier G, Stanojević R, Leith J, Shorten R (2007) A Critique of Recently Proposed Buffer-Sizing Strategies. ACM SIGCOMM Computer Communication Review 37(1):43–48

Wang Z, Yang X (2005) Birth-and-Death Processes and Markov Chains, 3rd edn. (40) Monographs in Pure and Applied Mathematics, Science Press, Beijing, in Chinese

Weiss H (1982) Economic Order Quantity Models with Nonlinear Holding Costs. Eur J Oper Res 9:56–60

Wischik D, McKeown N (2005) Part I: Buffer Sizes for Core Routers. ACM SIGCOMM CCR 35(2):75–78

Xu L, Harfoush K, Rhee I (2004) Binary Increase Congestion Control for Fast Long-Distance Networks. In: Proc. of IEEE INFOCOM

Xu Y, Chao X (2009) Dynamic Pricing and Inventory Control for a Production System with Average Profit Criterion. Probab Engrg Inform Sci 23:489–513

Yan H, Zhang J, Guo X (2008) Continuous-Time Markov Decision Processes with Unbounded Transition and Discounted-Reward Rates. Stoch Ana Appl 26(2):209–231

Yushkevich A (1977) Controlled Markov Models with Countable State Space and Continuous Time. Theory Probab Appl 22(2):215–235

Zhang J, Coyle E (1991) The Transient Solution of Time-Dependent M/M/1 Queues. IEEE Trans Inform Theory 37(6):1690–1696

Zhang Y, Loguinov D (2008) ABS: Adaptive Buffer Sizing for Heterogeneous Networks. In: Proc. of IEEE IWQoS

Zhang Y, Piunovskiy A, Ayesta U, Avrachenkov K (2010) Convergence of Trajectories and Optimal Buffer Sizing for MIMD Congestion Control. Com Com 33(2):149–159

Zheng Y, Zipkin P (1990) A Queueing Model to Analyze the Value of Centralized Inventory Information. Oper Res 38(2):296–307

Zhu Q, Prieto-Rumeau T (2008) Bias and Overtaking Optimality for Continuous-Time Jump Markov Decision Processes in Polish Spaces. J Appl Probab 45(2):417–429